This electronic thesis or dissertation has been downloaded from the King's Research Portal at https://kclpure.kcl.ac.uk/portal/

Phenomenal concepts and the Private Language Argument

Rodger, Jordan Michael

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT

Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. https://creativecommons.org/licenses/by-nc-nd/4.0/

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 06. May. 2023

Phenomenal concepts and the Private Language Argument

Jordan M. Rodger

Master of Philosophical Studies
King's College, London

Submitted November 2014

Abstract

This thesis makes the case for a Wittgensteinian account of phenomenal concepts, both in exegetical and justificatory terms. In terms of exegesis, I respond to critics who argue that Wittgenstein's so-called Private Language Argument (hereafter 'PLA') is incompatible with the existence of phenomenal concepts. In terms of justification, I offer a proposal for how to positively characterise phenomenal concepts in a way in which Wittgenstein might have endorsed. My proposal contrasts with other recent attempts to do so.

My argument is in four parts:

Part I introduces phenomenal concepts by way of their development out of materialist responses to Jackson's (1982) version of the Knowledge Argument. I provide my own attempted definition for how to understand the term 'phenomenal concept'.

Part II discusses Balog's (2009) and Papineau's (2002, 2011) criticism that the PLA is incompatible with phenomenal concepts. I reject this criticism on four grounds, arguing that it fails both as a criticism in itself, and as truly representing a proto-Wittgensteinian account of phenomenal concepts.

Part III discusses externalist interpretations of the PLA (Child, 2011, forthcoming, Hacker, 1993a, 1993b, Pears, 2008). I reject externalism on three grounds, where the uniting theme is the idea that we can make sense of phenomenal concepts in detachment from external circumstance without all loss of meaning.

Part IV is more positive. I make the case for a Wittgensteinian reading of phenomenal concepts where the crucial feature is 'the feeling/sensation itself'. It is not its relation to external counterparts. This makes room for an alternative reading of the point of the PLA. Under my reading, the 'logically private' sensation is not one that lacks an external counterpart, but one which lacks any connection to any other concepts. This investigation puts a different spin on just *who* and *what* Wittgenstein's target is in the PLA.

TABLE OF CONTENTS

Phenomenal concepts and the Private Language Argument

Précis

This essay is concerned with making the case for three chief claims:

(a) The criticism that Wittgenstein's so-called Private Language Argument (hereafter 'PLA') is incompatible with the existence of phenomenal concepts is to be rejected.

(b) The externalist attempt to positively define a Wittgensteinian approach to phenomenal concepts is to be rejected. I argue that externalism itself has serious flaws and also that Wittgenstein was not an externalist.

(c) An alternative proposal about the real target of the PLA is offered, specifying a new way of understanding what does and what does not count as *logically* private.

The structure of the argument runs as follows:

Part I introduces phenomenal concepts by way of their development out of materialist responses to Jackson's (1982) version of the Knowledge Argument. I also define how I will be employing the term 'phenomenal concept' throughout the essay.

Part II is concerned with providing a detailed response to the recent criticism of the PLA as presented by Balog (2009) and particularly Papineau (2002, 2011). The criticism argues that the PLA is incompatible with the existence of phenomenal concepts, and that this is 'bad news' for Wittgenstein because there are such concepts. This criticism is supposedly backed up by a version of Jackson's Knowledge Argument. This criticism is rejected on four fronts:

(i) I argue that the example used to support the Knowledge Argument is inadequate; Marianna's formation of a phenomenal concept is due to more than just acquaintance. Furthermore, if Wittgenstein is right, meaning *always* involves more than just acquaintance via ostensive definition.

(ii) Balog and Papineau are wrong in interpreting the PLA as expressing a 'public check' approach to concepts.

(iii) It is mistaken to think that the PLA is incompatible with phenomenal concepts in the first place; this is to conflate two senses of privacy. Wittgenstein is not an opponent of phenomenal concepts.

(iv) The previous responses are negatively oriented, they suggest why Wittgenstein might *not* be opposed to phenomenal concepts. But we can be more positive than this. Child (forthcoming) argues that a version of Knowledge Argument can be naturally developed from commitments Wittgenstein expresses elsewhere, suggesting that Wittgenstein would readily allow for phenomenal concepts. I agree with this view, and argue that if we permit a broader understanding of phenomenal concepts than is usual in the literature, then we can draw further textual evidence supporting a Wittgensteinian rendering of phenomenal concepts.

Part III discusses various externalist interpretations of the PLA (Child, 2011, forthcoming, Hacker, 1993a, 1993b, Pears, 2008). Child's argument is of particular interest in its explicit attempt to defend the view that phenomenal concepts must be 'tied to external circumstances, physiology, or behaviour' in order to be meaningful. I reject this externalism on three fronts. The uniting theme behind these objections is the idea that we can make sense of phenomenal concepts in detachment from external circumstance without thereby losing all meaning.

In Part IV, I suggest that the above objections to the externalist interpretation should not trouble the Wittgensteinian because Wittgenstein was not a committed externalist. In contrast to many interpretations, I argue that it is sensation S, *not* unexpressible pain sensations, which represents the first genuine instance of a *logically* private concept in *Philosophical Investigations*. This view suggests the PLA is aimed at a different target than is typically argued for. Moreover, it suggests a different understanding of what counts as *non-logically* private. Private sensations are not precluded by the community's inability to access them directly, or by their lack of connection with external behaviour. On the contrary, I argue that it is part of the 'grammar' of such concepts that their intrinsic features are, fundamentally, hidden to others.

PART I: Background and definitions

Before proceeding with the main arguments of this essay, it is necessary to provide some background and define some key terms.

The current debates surrounding phenomenal concepts have their origins in a set of responses to materialism, stretching back at least as far as the 1970s (Thomas Nagel's *What is it like to be a bat?* (1974) might be earmarked as the starting point for such responses). Materialism is here broadly defined as the view that the world's ontology, including the ontology of mental states, can be fully accounted for by a fundamental physics. This view has come under criticism from many quarters, but the overriding complaint is the same: Materialism, it is argued, is incapable of accounting for what has been variously termed as, e.g., the "what it is like' character', the 'phenomenal character', the 'subjective character', or the 'qualitative character' of consciousness:

> ...the fact that an organism has conscious experience *at all* means, basically, that there is something it is like to *be* that organism. ... fundamentally an organism has conscious mental states if and only if there is something it is like to *be* that organism — something it is like *for* the organism (Nagel, 1974, p.436)

We may describe the phenomenal character as the way that experience *feels* to the experiencer. There is 'something it is like' to taste a cool beer on a hot summers day, and this feels different to 'what it is like' to drink a hot coffee on a cold winters night. It feels different to see a bright red expanse as opposed to a pale blue one. And what it is like to be a bat is probably very different from what it is like to be Queen Elizabeth II. The important point to emphasise here is that for all these examples, their phenomenal character seems to reveal a property which is essential to consciousness; we couldn't have the consciousness without its possessing phenomenal character. There is no such thing as nonphenomenal forms of consciousness. It is the existence of this phenomenal character which, it is argued, prevents materialism from providing a full account of the world's ontology.

To be clear, the objection to materialism is not about whether material conditions play *a* part in explaining the existence of phenomenal character. On this point most materialists and non-materialists are in agreement.[1] Rather, the objection denies only that such explanation 'exhausts their analysis' (ibid. p.437). This thought is expressed by Joseph Levine (1983), who argues that any physical explanation of experience leaves an 'explanatory gap' between it and the phenomenal character of experience. That is, for all that we might have discovered about the workings of C-fibre firings in relation to the physical

[1] Barring eliminativist materialists, who deny that the term 'phenomenal character' has any place within an acceptable ontology (see e.g. Churchland (1986) and Stich (1983)).

bases of pain, 'there seems to be nothing about C-fiber firing which makes it naturally "fit" the phenomenal properties of pain, any more than it would fit some other set of phenomenal properties ... the connection between it and what we identify it with [is left] completely mysterious' (ibid. p.357). That is, C-fibre firing does little to explain *why* pain feels the way it does, instead of its feeling some other way, or no way at all.

In similar vein, Frank Jackson's (1982) version of the Knowledge Argument is used to support the conclusion that whilst someone might know all the physical facts about consciousness, they may yet be ignorant of *all* the facts about consciousness. Specifically, they might remain ignorant of the phenomenal character through which objects and experiences are represented and conceptualised in consciousness. To substantiate this argument, Jackson illustrates with the now familiar fictional example of Mary. Mary is a brilliant colour scientist, so brilliant that she comes to know all physical facts about colour vision. However, Mary has been imprisoned within a black and white room her entire life and so has never experienced colour herself. When she is finally released from her room and presented with her first non-black or white colour, let's say it is a red rose, Mary seems to learn something new about the world, specifically about its visual appearance. She learns that the visual appearance of red consists of a particular kind of phenomenal property, a property of which she was ignorant before her release. She learns 'what it is like' to see red. Accepting that complete physical knowledge does not mean that Mary has complete knowledge, it follows that physical facts are not the only facts. Therefore, materialism is false. Or so the argument runs.

Whilst defining phenomenal character in any detail has proved divisive,[2] there has reached a certain level of consensus over some key issues. There is near consensus that the Knowledge Argument is successful in demonstrating, at a minimum, that Mary attains new knowledge after her first acquaintance with new visual experiences.[3] Agreement has also been largely reached regarding the premise that this new knowledge counts as genuinely *propositional* in form, over and above any new *practical ability* Mary might acquire.[4] A significant proportion of discussion has now led to asking whether this new propositional knowledge counts as new knowledge at the level of *reference* or at the level of *concepts (*or *sense*). The term 'phenomenal concept' arose out of an attempt by some materialists to suggest that Mary's new knowledge is restricted to the conceptual level, and does not interfere with the metaphysical (see e.g. Loar, 1990, Lycan, 1990, Tye, 1995, Perry, 2001, Papineau, 2002, chapter 2). Relating to Mary, this materialist response accepts that Mary learns something new when she sees red for the first time, but denies that this

[2] Crane (2000) has persuasively argued that the term 'qualia' (essentially what I term as 'phenomenal character') lacks a uniform meaning. Sometimes, it refers to supposed properties of the objects of experience (i.e. sense-data), sometimes to supposed properties of experiences themselves, and often the use equivocates between the two. Unfortunately, I cannot go into this argument in more detail.

[3] Although see Dennett (2007) who presents an interesting argument to suggest that if Mary has all physical knowledge, then she will already know all there is to know prior to exiting.

[4] Although see Nemirow (2006) for continuing disagreement.

new knowledge requires the recognition of a distinct ontological property. Just as we learn something new when we learn for the first time that Clark Kent is, in fact, Superman, so too Mary learns something new when she sees red for the first time. In both cases, the subjects gain a new way of thinking about Superman or about redness, and in this sense gain new concepts. But this doesn't mean they are now referring to an additional property. Rather, they are referring to the *same (material) property* as before, but in a new way, under a new guise.

One professed advantage of this version of materialism (a 'Type-B materialism' (Chalmers, 1996) or 'Phenomenal Concept Strategy' approach (Stoljar, 2005)) is its frank admission that phenomenal character cannot be derived a priori from physical or functional explanation. Instead, the explanation is determined a posteriori; the identity is discovered as a result of scientifically-informed empirical investigation. So we *should* be surprised at how it turned out that phenomenal character *is* identical with being in a given physical or functional state. There are no conceptual entailments between neurobiological descriptions of brain states and phenomenal character. And so it is quite possible to conceive phenomenal character without brain activity and brain activity without phenomenal character. The Type-B materialist can seemingly accept that Mary makes genuine epistemic progress, whilst also retaining the intuition that the physical world is exhaustive of what there is. For the Type-B materialist, Mary's progress is *conceptual*, not ontological. Most commentators now accept that there are such phenomenal concepts.[5] The shift of attention has now turned to whether phenomenal concepts do in fact lead us towards such materialist conclusions, or whether, as Chalmers, puts it: 'no account of phenomenal concepts is both powerful enough to explain our epistemic situation with regard to consciousness and tame enough to be explained in physical terms' (2010, p.306, also see ibid. chapter 10).

Although the term phenomenal concept initially arose out of these metaphysical debates, the metaphysics of phenomenal concepts is not of direct concern for my purposes in this essay. Insomuch as it is possible, I therefore remain neutral on the metaphysics.[6] My interest instead revolves around what logical and epistemological conditions phenomenal concepts must satisfy in order to count as genuine concepts, and whether or not they do in fact satisfy them.

To start, it is helpful to specify phenomenal concepts in at least their broad outline. Daniel Stoljar succinctly describes a phenomenal concept as:

> ...a concept of a specific type of perceptual or sensory experience where the notion of experience is understood phenomenologically (2005, p.269)

[5] Although see Ball (2009) and Tye (2009).
[6] To discuss the metaphysics in any detail would probably amount to double the present word count.

Whilst this statement certainly requires elaboration, it does provide a neat summary: a phenomenal concept is a concept of the *phenomenological* character of the experience. Phenomenal concepts characterise the specifically qualitative, subjective, *phenomenal* features of experience. This much is clear. However, similarly to discussion about phenomenal character, many proposals have been offered to define the term 'phenomenal concept' in more detail. And, again, when it comes to the specifics of exposition, the details lead to a variety of divergent positions. Most accounts of phenomenal concepts say that it includes features such as that the concept is essentially first-personal, whose content is directly accessible only to the subject. It is accepted that phenomenal concepts are acquired (typically) through acquaintance with the experience. Most accept that the experience is often of finer grain than the concept, and that the concept is often incorrigibly applied when used to directly refer to ones present experiences (see Balog, 2009, pp.299-300 for a more comprehensive breakdown). More controversially, phenomenal concepts are variously regarded as recognitional concepts (Loar, 1990), indexical concepts (Perry, 2001), perceptual concepts (Papineau, 2002, 2006), sui generis concepts (Chalmers, 2003).

With slim hope for providing an all-encompassing definition of this term, I will not pretend to do so. In its place is my own way of formulating matters. I take a broadly Wittgensteinian starting point by suggesting that a complete theory of concepts (of concepts *in general*, not just phenomenal concepts) must be capable of providing an accurate account of at least the following two conditions:

Formation conditions: The conditions under which we can be said to acquire the concept, i.e. how we came by that concept in the first place. (See e.g. *Philosophical Investigations* §§1-33)

Application conditions: The conditions under which employment of that concept is rightly or wrongly determined; i.e. what constraints govern the use of the concept. (See e.g. *Philosophical Investigations* §§185-242)

According to these conditions, one must be able (in principle) to explain how acquisition of the prospective concept is achieved, and how the concept gets used correctly in practice. The suggestion is that these conditions provide a sensible platform from which to understand particular instances of the supposed concept.[7]

[7] In thinking about why my reader should be persuaded to accept this way of formulating matters, I cannot in all honesty provide a particularly clear answer. I was then pleased to read a neat passage in Peacocke's *A Study of Concepts* (1992):

> Hume, Kant, Wittgenstein, early and late, Quine, and recent realists and antirealists have all carried through projects of that general description. While agreeing on little else, all these thinkers are committed to holding that without a general treatment of concepts one will not have a satisfying philosophy ... The general form sets a standard to which accounts of particular concepts must conform. Insofar as certain alleged concepts are declared to be spurious, as beyond the limits of

To apply these conditions to task, then, we might ask what constitutes the formation conditions and application conditions governing phenomenal concepts. One response might be as follows: The distinctness of phenomenal concepts is partly a result of the fact that they are (typically) available only to those who have been directly acquainted with the requisite phenomenal experience(s). It is only acquaintance with that experience which determines consequent use—in virtue of having the experience, one can use and misuse the concept. In short:

PC formation conditions: Acquaintance with experience is (typically) necessary for forming the concept.

PC application conditions: Rules for use are determined solely in virtue of that acquaintance.

In this manner *PC application conditions* are dependent on *PC formation conditions*. Part II will introduce a more detailed discussion of these conditions, and pose the question as to whether such conditions are in violation of a Wittgensteinian approach to concepts, or whether there is an acceptably Wittgensteinian presentation of phenomenal concepts in light of them.

To make two final asides; the term '(typically)' will be frequently expressed in the initial part of this essay (as it has been so far) to indicate a caveat that is intended throughout. The caveat is this: The possession of a phenomenal concept will *typically* be acquired by means of one's having first been acquainted with the appropriate phenomenal particular(s). As some philosophers draw attention to, whilst this is *in fact* how phenomenal concepts are acquired, it does not mean that phenomenal concepts are *of necessity* acquired in that way. According to Conee (1994, p.140) and Peacocke (2009, p.188), one may grasp a phenomenal concept without having actually experienced the phenomenal character oneself—one may *interpolate*, as presumably Hume's imaginer of the missing shade of blue would do. Or, more bizarrely, one might be duplicated, memories and all, and so have the phenomenal concept without having had the experience itself (see Dennett's example of Swamp Mary (2007, pp.24-5)). Even so, the vast majority of instances of phenomenal concepts arise in this way; acquaintance with the experience is thus the 'typical' way phenomenal concepts are acquired.

Final aside; Wittgenstein never used the term 'phenomenal concept', the term was not invented until recently. The closest term (arguably) is 'sensation concept', at least if we are considering the way in which 'phenomenal concepts' are typically understood in the literature. For the purposes of this essay, the terms

---

> genuine thought, it must also be by appeal to properties of this general form that such claims are justified. (p.vi)

This passage chimes with my own sentiments about methodological priority.

are being treated synonymously, except for when discussing examples where the difference is indicated. With these asides in mind, we move on to Part II.

PART II: Wittgenstein, opponent of phenomenal concepts?

The Private Language Argument

The term 'Private Language Argument' was not used by Wittgenstein, and it would be mistaken to think that there is just one argument provided in the relevant sections of *Philosophical Investigations*, §§243-315.[8] As one commentator puts it: 'the texts in which the [PLA] is developed present a running battle against Private Language rather than a single engagement' (Pears, 2006, p.37). Indeed, a brief reading of these sections shows that Wittgenstein not only presents *many* arguments against the possibility of a private language (see e.g. Schroeder, 2006, pp.201-215), but that a whole range of other philosophical issues are introduced as well. The 'argument' does not extend from premises to conclusion in linear form, as per the typical requirements of modern academic philosophy. Rather, the sections proceed via a barrage of strange thought-experiments. The reader is presented with a variety of situations where language and concepts function in ways very different to our own, where the environment itself is somehow different, where there are talking pots, stones that feel pain, babies that feign pleasure, and of course we are introduced to the infamous speaker of the private language.[9] We are repeatedly encouraged to compare and contrast these language-games to our own (as 'objects of comparison' (*PI* §130)). Conclusions are rarely, if ever, explicitly drawn, or the thoughts summarised to help the reader collate the ideas. For better or worse, this was Wittgenstein's preferred style:

> [M]y thoughts were soon crippled if I tried to force them on in any single direction against their natural inclination.—And this was, of course, connected with the very nature of the investigation. For this compels us to travel over a wide field of thought criss-cross in every direction (*PI*, *preface ix*)

Of course, in the same passage Wittgenstein also admits to 'not liking [his] writing to spare other people the trouble of thinking' (*PI, preface x*[e]). Given this 'wide field of thought', it is inevitable that commentators

[8] Hereafter, I use '*PI*' when referring to sections in *Philosophical Investigations*. I use similar acronyms for references to Wittgenstein's work, as set out in the bibliography.

[9] From these sections, it should come as no surprise to learn that Wittgenstein enjoyed reading Lewis Carroll (see Pitcher (1965).

pick up on different strands of thinking, emphasise different themes, and provide very different reconstructions regarding the ultimate points and purposes of the PLA.

Fortunately, the introduction of the PLA is a somewhat more straightforward matter:

> A human being can encourage himself, give himself orders, obey, blame and punish himself; he can ask himself a question and answer it. We could even imagine human beings who spoke only in monologue; who accompanied their activities by talking to themselves.—An explorer who watched them and listened to their talk might succeed in translating their language into ours. ...
>
> But could we also imagine a language in which a person could write down or give vocal expression to his inner experiences—his feelings, moods, and the rest—for his private use?—Well, can't we do so in our ordinary language?—But that is not what I mean. The individual words of this language are to refer to what can only be known to the person speaking; to his immediate private sensations. So another person cannot understand the language. (*PI* §243)

We are first asked to consider cases where we talk to ourselves in different ways. A relatively humdrum thought experiment. Next is to imagine a community where people *only* talk to themselves; slightly more bizarre, but not beyond the realms of imagination (after all, the imagined explorer can do it). Finally, we are asked to think about a quite different question: Can we imagine a language where a person can write or talk about their *private inner experiences*? This is the fundamental question of the passage, and indeed of the PLA as a whole. It should not be understated. However, just what this question is getting at only becomes more apparent from the ensuing dialogue between Wittgenstein and his assumed interlocutor. The interlocutor is puzzled, thinking that Wittgenstein is asking him to imagine something which we actually do all the time— namely, write down or talk about our inner feelings and moods. The interlocutor assumes that Wittgenstein is restating something similar to the first question posed. Wittgenstein's response to this puzzlement is central to the PLA; it is here that the PLA is clarified and defined. The interlocutor's assumption is *not* what Wittgenstein means at all. Rather, the language is imaginary, where the 'individual words of this language are to refer to what can *only* be known to the person speaking; to his immediate private sensations. So another person *cannot* understand the language' (emphasis added). The language is not one like ours, because in this imagined language the words refer *only* to immediate private sensations, and therefore words whose meaning is *in principle*, *logically*, un-understandable by others.

Unfortunately (or 'interestingly' depending on your persuasion), just what *this* all means, Wittgenstein never quite seems to spell out. In one sense, the target of this section is relatively clear; it is asking whether there can be a *private* sensation language, a language which refers to *private* sensations, sensations which *cannot* be understood by anybody else. This purported language is set in contrast with our ordinary language. And the following sections revolve around the discussion and rejection of the idea that there could be such a language. However, in another sense, we are instead taken deeper into the rabbit hole. For we are led to ask: What, then, could these logically private inner experiences/ sensations be? What is a private language? What is it that makes this language different from ours? What is wrong with it? Who thinks that there is one? What is the philosophical significance of these reflections? The content of the PLA revolves around the dialogue between Wittgenstein and his interlocutor over just these kinds of questions. Commentators have disagreed ever since.

The bulk of the remainder of this essay is concerned with discussing those interpretations of the PLA which have been canvassed in view of the recently discussed topic of phenomenal concepts. Specifically, I respond to the objection that the PLA is incompatible with phenomenal concepts (this makes up Part II), and suggest in what way the PLA *is* compatible with such concepts (Parts III and IV).

The Objection

Katalin Balog (2009) and David Papineau (2002, 2011) have recently criticised the PLA for its alleged incompatibility with the existence of phenomenal concepts. The objection can be stated in three relatively straightforward steps. In the first step, it is suggested that the point of the PLA is to show that concepts must meet certain constraints in relation to their possibility for public scrutiny prior to meaningful employment. (The 'priority' claim here is stated explicitly by Papineau (see his 2002, p.129, 2011, pp.181-2). As per Balog:

> Wittgenstein, in his famous private language argument, argues that for a term (concept) to have meaning (or reference) it must be possible to intersubjectively check whether an application of that term is correct (2009, p.298)

As per Papineau:

> ...the main point of the private language argument is to make it clear that private terms ... must refer to something publically accessible if they are to be genuinely meaningful (2011, p.182)

In the second step, it is argued to be an essential feature of phenomenal concepts that their acquisition and exercise occurs strictly within the private domain: the subject 'breathe[s] meaning into her term purely by focusing inwards and attaching it to her current experience' (ibid. p.181). So it is (typically) through acquaintance with the experience that determines the meaning of the concept, and only thereafter that one ascribes that meaning publically.

On the basis of these first two steps, the proposed incompatibility between the PLA and phenomenal concepts is evident: the PLA states that the acquisition and exercise of concepts must be susceptible to public scrutiny. Phenomenal concepts are purportedly formed and exercised independently of such scrutiny.

At the third step, Papineau argues that a successful variant of Jackson's Knowledge Argument, as involving Marianna (Nida-Rümelin (1996)), proves there to be distinct phenomenal concepts. In conclusion, the PLA is not only incompatible with phenomenal concepts, but such incompatibility is ultimately 'bad for Wittgenstein' (Papineau, 2011, p.175). Or so the objection runs.

Given its centrality to the argument, it is worth considering the Marianna example in more detail. In the case of Mary, there is some ambiguity about how to explain what Mary believes when she is released from her room and thinks "Ah, so that's a *red* rose." To clarify; suppose Mary is unwittingly presented with an imitation *blue* rose, under the pretence that it is red. Under one reading, Mary believes that the rose is blue-coloured—she sees the blue-coloured rose, assumes her colour experiences are similar to others, and so forms the belief that roses are that colour, i.e. *blue*. However, this is problematic because if Mary is asked what colour the rose is, she will surely say that it is a *red*-coloured rose; so we might also be inclined to say Mary believes she sees a red-coloured rose. It is not immediately clear that Mary learns 'what it is like to see red' simply in virtue of exiting the room, even in cases where there is no such trickery involved.[10]

Nida-Rümelin (1996, pp.221-5) introduces an example which aims to dispel the ambiguity surrounding Mary's belief by distilling the singularly phenomenal element of what Mary believes apart from anything which may confuse issues. Unlike Mary, Marianna is shown a coloured piece of paper and isn't told the name of the colour she is looking at. So Marianna isn't able to infer the name of the colour on the basis of its shape, or the sort of object she is presented with. Marianna is not expected to form any beliefs regarding the name of the colour, beyond those arrived at through mere guesswork. And yet we can

[10] We might envisage the subject landing herself in similar sorts of puzzles as Kripke's Pierre, seemingly both liking and disliking the colour red (see Kripke, 1979).

envisage Marianna coining a concept to refer to the colour of the experience that she just saw. Moreover, she is not only capable of coining a concept, 'but can classify it visually, and can re-create it in visual imagination' (Papineau, 2002, p.108, also see ibid. p.128 and his 2011, pp.176-7). That is, she can now perform such cognitive tasks as recognise new particular instances of it when she sees it again, and she can imagine it being the colour of the sky, or the colour of the sun, and so forth.[11] None of which she could do before her encounter.

These considerations lead Papineau to state why Marianna acts as a counter-example to the PLA:

> Marianna isn't just thinking about subjective states with the help of her imaginative and introspective powers. She coined a special term Φ to refer to a type of experience. And this term definitely seems to have two features that Wittgenstein took to discredit the idea of a 'private' language: first, Marianna is supposed to have breathed meaning into her term purely by focusing inwards and attaching it to her current experience; second, the term so formed will be one whose meaning will be incommunicable to anyone except Marianna herself. (2011, pp.180-181)

Crucially, Marianna is able to both acquire and exercise such concepts *solely in virtue of being acquainted with the experience*. Furthermore, Marianna will be capable of forming this concept even though she is incapable of communicating to anyone else about the experience her newly coined concept refers to (see Papineau, 2002, pp.127-132, and 2011, p.183). And this stands in contradiction to the PLA. Or so the argument runs.

Responses

In what follows, I reject Papineau's claim about the following two purportedly anti-Wittgensteinian features of Marianna's situation:

(i) Marianna's ability to 'breathe meaning' into a concept solely through acquaintance.
(ii) Mariana's ability to do so without being able to communicate what she means to others.

The main point of objection is that Papineau is wrong to think Marianna's situation represents an objection to a Wittgensteinian approach to concepts. In my view, Wittgenstein is no opponent of

[11] Unlike the puzzle of Kripke's Pierre, which (as per the previous footnote) troubles Mary so, Marianna might evade these troubles insofar as her belief is squarely attached to *that* experience, whatever it is called.

phenomenal concepts and Marianna presents no objection to a Wittgensteinian understanding of such concepts. In what follows, I give three arguments in support of this claim:

(a) Wittgenstein persuasively rejects the idea that a concept can be formed through 'acquaintance alone.' So the Marianna counter-example fails if it is used to defend this idea. I suggest that this need not lead to a rejection of phenomenal concepts outright, but rather points towards the need for a revision of how we understand their origins.

(b) I briefly restate the common response to a Kripkean interpretation of Wittgenstein's 'public check' or 'community' view of rule-following.

(c) I reject the idea that Wittgenstein would object to the kind of incommunicability that Marianna is faced with; Marianna's privacy is contingent and therefore not the target of the PLA.

If the dialectic of argument is on track, there is so far no reason why Wittgenstein cannot account for the existence of phenomenal concepts. The subject of Parts III and IV will go on to suggest what kind of privacy Wittgenstein *is* objecting to.

Response 1:

In a lecture delivered in 1934, Wittgenstein gives an example of a theory of meaning opposed to his own:

> The following case seems to contradict the claim that the use of a word does not follow from the idea: by an example, i.e., by an ostensive definition, we are able to give a person an idea of *red,* say. We show him the *meaning* of the word "red". If we can give the meaning by ostensive definition, then the correct use will follow from its meaning and not from the rules. The correct use of the word "red" is thought of as a consequence of its meaning, which is given in one act, all at once. (*AWL,* p.87)

The central idea is clearly put: correct use will *follow from* the meaning. The idea is recognisably anti-Wittgensteinian (or *Augustinian*, *PI* §§1-33) in its attempt to divorce meaning from use. In this context, 'meaning' must be intended in the sense of 'bedeutung' and so more closely with the English word 'reference'. The *bedeutung* is provided as a result of successful ostensive definition. Ostensive definition is itself defined as the act of explaining the meaning of a word, typically by successfully pointing out,

ostending, the relevant object or experience (i.e. the referent) to the receptive observer. In the example above, it is in virtue of having an example of red pointed out that one acquires the concept *red*, and thereby (*following from*) learns how to use it correctly. In slogan-form: Use follows meaning.

Papineau defends Marianna as representing a case in favour of just this kind of theory:

> Since [Marianna] clearly can think good thoughts with her new concept, say I, it follows that normative rules are inessential to representational content ... Content does not derive from normative rules, but rather from the kind of non-normative natural facts invoked by causal or teleosemantic theories of representation. In so far as there are norms in the area of judgement, these *follow* from the prior constitution of content, and are not a precondition thereof ... it is no deficiency in [Marianna's] concept that she is not sensitive to any normative principles tying its use to public criteria (2002, pp.128-9. Also see his 1999).

My aim in what follows is to defend Wittgenstein by suggesting why Marianna does not count as a good example of someone who is 'not sensitive to any normative principles'. On the contrary, Marianna is well primed to think those kinds of thoughts precisely because she *is* a part of the right normative environment.

Wittgenstein's immediate response to the acquaintance-based view in the lecture quoted above is as follows:

> *However*, note that the use of the word is not actually fixed by giving someone, by ostensive definition, what is supposedly the meaning. For he may now use "red" when he sees a square. (*AWL*, p.87)

Wittgenstein's quipped response here seems to be that even in a supposedly straightforward case where ostensive definition is used to determine the concept 'red', it is possible to imagine that the subject fails to do so, or does so in unusual fashion. The point is not (*pace* Kripke (1982, see response 2 below)) a sceptical one, namely that 'there can be no such thing as meaning anything by any word' (ibid. p.55). Rather, the point seems to be that, so far as we are limited to thinking about the sequence of acquaintance contact, there is nothing to suppose that the ostensive definition means the shape of the object instead of the colour, or perhaps only red-coloured squares. We might even imagine the problems beginning further back, where one may be perennially disposed to follow the tip of the finger up towards the arm, not from arm to fingertip, like a cat (*PI* §185). Again, the main point is that there is nothing in the act of ostensive definition that *of itself* entails meaning will be secured, or secured in the way we expect. As famously phrased in *Philosophical Investigations*:

> An ostensive definition can be variously interpreted in *every* case. *PI*§28

In another lecture, Wittgenstein develops this thought:

> If I showed a person a red sample and asked him to bring me something red, he would do so. But he need not. He might bring me something having the complementary colour, though this is not usual. One use is not more direct than another—only more usual. We are extraordinarily affected by the way in which we do in fact react to a sign. The result is that certain ideas stand to us for certain uses because that is how we usually apply them. We therefore think that those ideas have that most usual use *in* them, though they could perfectly well be imagined to have another use (*AWL*, pp.88-9)

I take these remarks to be expressing, at least in part, the idea that a philosophical understanding of meaning, of understanding how people come to know how to use concepts and so forth, must involve consideration of features beyond the act of acquaintance. That is, it must involve recognition of the deeper web of supporting features which contribute towards explaining the role of concepts in our lives. Even the act of pointing out or gesturing towards objects to signify what a concept refers to is itself an action saturated with normative significance, not an independent explanation of it (*PI* §16). As such, if we wish to explain how meaning arises, we need to extend beyond the act itself, and situate it in its broader normative context:

> What determines our judgement, our concepts and reactions, is not what *one* man is doing *now*, an individual action, but the whole hurly-burly of human actions, the background against which we see any action. *Z* §567

This approach rejects the idea that we can explain meaning by beginning with our acquaintance with objects or experiences and then branching out, but from the fact that we are situated in a normative environment which is amenable to such acquisition and application in the first place. These considerations are not used to deny the role of ostensive definition in practice, or deny that one can go on to use a concept in virtue of being shown a sample. Acquaintance is important ((typically) *necessary*), but only when it is situated in the right environment: 'an ostensive definition explains the use—the meaning—of a word when the overall role of the word in language is clear' (*PI* §30, also see *PI* §§31-32). The criticism, then, is squarely that the acquaintance-based view fails to pay heed to the complexity of the situation (see *LCA*, p.2).

Dialectically speaking, we reach loggerheads. Papineau says acquaintance explains concept acquisition, Wittgenstein says it doesn't. One way to move the argument forward is to specify an example where the conditions for ostensive definition are in place, but where the 'normative rules' are altered or, as much as possible, removed from the scenario altogether. I take such an example to be a continuance of the Mary thought-experiments: Just as Papineau objects to the original Mary thought-experiment for failing to disentangle alternative explanations about the content of Mary's belief, we might levy similar charges against Marianna. If the acquaintance-based view is right, if 'normative rules are inessential to representational content', then Marianna's normative education aids only to confuse the essential facts about her concept acquisition, which is achieved through 'acquaintance alone.' And if this is so, then we should be able to reconceive changes to Marianna's education, but without affecting those essential facts. That is, if Marianna can be shown to form a new concept under such conditions, then Papineau's conclusion that 'acquaintance alone' suffices to elicit the formation of a new concept gains substantial plausibility. However, if such changes to the normative rules impact negatively on Marianna's ability to form such concepts, then this indicates that Marianna never 'breathed meaning into her term purely by focusing inwards and attaching it to her current experience'; there is *always* more to say.[12]

Before reimagining the Marianna example, it is important to highlight those factors which, from a Wittgensteinian perspective, seem to make Marianna a good candidate for forming phenomenal concepts, but which have little to do with the overt act of ostensive definition. There are features such as the fact that Marianna possesses a colour vocabulary, and is aware that certain objects are coloured in ways that that vocabulary aims to represent. She is aware of which objects are typically associated with which colours, by name at least. There is the fact that colours themselves are relatively stable, although subject to some variation through differing lighting conditions, and Marianna is aware of this. Marianna has also grown up with an education and in a culture which pays significant attention to colours. She will know that seeing a coloured piece of paper will give her no indication about the phenomenal character of the colour she sees. We might even imagine that Marianna has read Jackson's (1982) paper, and foresees the kind of conceptual difficulties she will face upon exiting. The premise I wish to defend is that these factors, and undoubtedly many more, can be regarded as influencing what happens to Marianna, conceptually-speaking, when she sees that coloured card for the first time. Such features leave Marianna primed ready to expect *something* to happen upon her exit.

To start thinking about what might happen to Marianna in a world with different normative rules, I appeal to a list of examples provided by Wittgenstein:

> In a world different from ours colours might play a different role. Think of various cases.

[12] Of course, proponents of the Knowledge Argument need not suggest that acquaintance alone suffices. This does not alter the fact that they, in fact, do suggest this.

(1) Certain colours are tied to certain forms. Circular shapes, red, rectangular ones, green, etc.
(2) Dyes can't be produced. You can't colour things.
(3) *One* colour always linked together with a foul smell, or poisonousness.
(4) A far greater incidence of colour-blindness than now exists.
(5) Different shades of grey abound; all other colours are extremely rare.
(6) We can reproduce a great many shades of colour from memory. If our number system is connected with the number of our fingers, then why shouldn't our system of colours be connected with the specific ways in which they occur.
(7) A colour occurs only in gradual transition into another one.
(8) Colours always occur in the sequence of colours in the rainbow.

RPP, *Vol. 2*, §658

My initial response to this list is to try and imagine each bizarre scenario. I am not even sure how to begin envisaging (8), and do not much understand the second part of (6). The overall impression the list makes, it seems, is someway clearer. If the world were any of these kinds of ways, then it is likely that our concepts, especially our colour concepts, would have very different functions compared to those that they have in our lives. Some effects may be relatively slight—(2), perhaps, or (3)—some much more radical. I find (5) particularly intriguing, it conjures up ideas of a society with conspiracy tales of their occurrence, or where religious significance may be attached to them, and so on. The list provides a springboard for thinking about how alterations to the physical constitution of the world and alterations to our cognitive and sensory constitutions might fundamentally affect the kinds of normative rules governing the use of colour concepts. The deeper message behind this list, perhaps, is to encourage us to think about the particularity, and the fragility, of the circumstances which give rise to our own uptake of concepts.

In a more detailed example, Wittgenstein considers another possibility:

> Suppose I were to come to a country where the colour of things—as I would say—changed constantly, say because of a peculiarity of the atmosphere. The inhabitants never see unchanging colours. Their grass looks green one moment, red at the next, etc. Could these people teach their children the words for colours?—First of all, it might be that their language *lacked* words for colours. And if we found this out we might explain it by saying that they had little or no use for certain language-games.
> *RPP,* Vol. 2, §198

Let us assume Marianna has been brought up in this environment. Let us further assume that the community does lack words for colours, due to their constant radical shifting. Furthermore, suppose that

the community lacks any other recognised way of referring to such colour shifts. In other words, the whole notion of colour is entirely foreign to this community, and plays no role in their practices. The relevant question is whether phenomenal concepts might survive in the face of these whole-sale changes.

Suppose we (as visitors to the community) attempt to replicate the coloured card experiment in the colour-changing world. We show a coloured card to Marianna, only this time noticing that it changes colour during the course of our showing. Let us suppose further that at some specified time during our showing, the card flashes a distinctive red colour. At this specified time, Marianna's situation in regards to the ostensive definition of coloured card is the same as the original case involving Marianna: She is situated in front of the coloured piece of card and, for that moment at least, her gaze is directed onto the sample of red in the same way. Moreover, our intention is to point out to Marianna the sample of red, we may even attempt to guide her by saying "That, Marianna, is what we mean by 'red'". Can we say, in this instance, that Marianna gains a new concept in virtue of the acquaintance?

If the only information Marianna is given about the colour red stems from what we do at that instant, then it seems to me more likely that Marianna will form no such concept of red. The reason for this is that there is slim evidence to suggest that she registers the colour in any way. She has no language to speak about it, it is doubtful whether she will visually classify the colour or imaginatively recreate it. Marianna's reaction will more likely be to question whether we mean the card itself by calling it 'red'. The whole point of the exercise, it seems, will pass Marianna by. Only if we then go on to call it something else when its colour has changed a few times may Marianna begin to get an inkling of what we are talking about. But then this takes us beyond the confines of ostensive definition; only once Marianna begins to get the right kind of 'training' (*PI* §§5-6) in what we mean, and some background about what we are trying to classify, will she start forming concepts. The first instance of ostensive definition, I argue, need not present Marianna with any special concept simply in virtue of her acquaintance:

> Do not believe that you have the concept of colour within you because you look at a coloured object—however you look. *Z* §332

This suggests that more is required than an act of ostensive definition, more than just 'acquaintance with an experience'.

In fact, we do not need to consider alterations to the physical world in order to imagine changes to Marianna's conceptual uptake. Equally relevant are Marianna's own 'interests' and the interests of her community. We can equally well imagine that she is part of a community that simply fails to recognise colours:

> I want to say: an education quite different from ours might also be the foundation for quite different concepts ... For here life would run on differently.—What interests us would not interest *them*. *Z* §§387-388

> Concepts lead us to make investigations; are the expressions of our interest, and direct our interest. *PI* §570

What such considerations bring out is just how much we rely on other factors beyond acquaintance in establishing meaning. As such, we are drawn towards the conclusion that in addition to acquaintance with the requisite experience, it is equally a prerequisite for the experiencing subject to be grounded within a normative framework that enables them to recognise and categorise such experiences. Without the right environmental conditions, without the right kind of 'training', without the right interests, there is no guarantee that acquaintance will successfully elicit phenomenal concepts.

Again and again, Wittgenstein uses examples and thought-experiments to emphasise *both similarities and dissimilarities* when compared with how our language functions (see *PI* §§130-1). The point of this exercise is intended to shake us out of seeing language function in one way: "In philosophy one feels *forced* to look at a concept in a certain way. What I do is to suggest, or even invent, other ways of looking at it" (as quoted in Malcolm, 1984, p.43). By altering the supporting conditions wherein phenomenal concepts are usually formed, and imagining potential knock-on effects, we are encouraged to consider the hotchpotch coming together of language, concepts, environment, physiology, and culture, and how they all interact in complex ways to contribute towards explaining how concept acquisition takes place. A general theory of 'acquaintance alone' simply fails to recognise this. In view of this, the claim that 'acquaintance' plays some special role is to be resisted, in favour of a more holistic attention to the surrounding context.

Do these considerations lead to the conclusion that Wittgenstein would object to the existence of phenomenal concepts? My answer to this is that if phenomenal concepts are defined as concepts of experiences that are formed solely through the acquaintance, then yes, Wittgenstein does oppose them. However if, as I wish to defend for the rest of this essay, we can specify an alternative account for how phenomenal concepts are acquired, one that respects the considerations alluded to above, then it is by no means clear that Wittgenstein is an opponent of phenomenal concepts. Rather, he is opposing only how some philosophers set about characterising them, and questions the plausibility of the idea that we can imagine someone who forms concepts irrespective of normative constraints of one kind or another. As will be argued for in Part IV, it is just this conception of concepts which is the real target of the PLA, where a *logically* private language is to be understood as a language which is devoid of all such normative ties.

Response 2:

Whatever we might conclude about its argumentative merits, Balog's and Papineau's criticism can hardly be regarded as adhering to the highest standards of exegetical rigour. Indeed, only two remarks from *Philosophical Investigations*, §§257-8, are referenced to defend their interpretations. Admittedly, exegetical rigour is not the object of their investigations. However if we want to know why we should believe that Wittgenstein defends the ideas they suggest, it is pertinent to ask how they justify attributing such views to Wittgenstein in the first place. Fortunately, insofar as they attribute Wittgenstein as advocating a 'public check' theory of meaning, their exegetical allegiances can be more easily surmised. The attribution of a public check theory of meaning chiefly derives from the 'Kripke-Fogelin' interpretation of the PLA (hereafter 'K-F interpretation').[13] According to the K-F interpretation, the idea that mental concepts must be susceptible to public scrutiny is regarded as an extension of the more general premise set out earlier in the *Investigations* (*PI* §§185-242), namely that *all* concepts must be susceptible to such checks.

The public check theory of meaning itself supposedly arises in response to a paradox that faces traditional referentialist theories of meaning:

> This was our paradox: no course of action could be determined by a rule, because any course of action can be made out to accord with the rule. *PI* §201

The K-F interpretation claims that Wittgenstein fully accepted the radical conclusion of this paradox, that 'there can be no such thing as meaning anything by any word' (Kripke, 1982, p.55), and offered a 'sceptical solution' in light of it. The solution was to adopt a 'Community Thesis' which recognises the rule-governing community as sole determinant of whether particular instances of concept application are successful in conveying meaning (see Kripke, 1982, chapters 2 and 3, Fogelin, 1987, chapter 11 and 12, especially pp.179-183).[14]

---

[13] I call this the 'K-F' interpretation because both Kripke and Fogelin were early proponents of the view that Wittgenstein defends a 'public checkability' theory of meaning. Fogelin even goes so far as to acknowledge that there is 'nothing important that distinguishes our two interpretations' (1987, p.242, fn.8).

[14] Other proponents of a version of Community Thesis include e.g. Norman Malcolm and John Canfield (see e.g. Malcolm, 1986, 1989, Canfield, 1996)). Not all versions of Community Thesis adopt Kripke and Fogelin's version of it, and there is certainly no a priori requirement to endorse the even more specific Public Check Argument either. For the purposes of this essay I have no determinate opinion over the wider issue of whether some version of Community Thesis might be correct. My scope is here limited to criticising the Kripke-Fogelin line.

One consequence of this move to the Community Thesis is supposedly summarised at *PI* §202:

> And hence also 'obeying a rule' is a practice. And to *think* one is obeying a rule is not to obey a rule. Hence it is not possible to obey a rule 'privately': otherwise thinking one was obeying a rule would be the same thing as obeying it.

In short, if the community determines correct and incorrect applications of rules, then there can be no such thing as a privately-determined rule. The reason for this is that without independent jurisdiction, whatever one decided to count as a correct application of the concept would therefore, in virtue of the very decision, *be* a correct application. And such a conclusion is surely awry, because there needs to be a distinction between something *seeming* to be a correct application and something *actually being* a correct application. This distinction cannot be maintained without an independent arbiter, and so it follows that public standards of checkability are a necessary condition for meaning (see e.g. Kripke, 1982, p.167).

The K-F interpretation then claims that the consequences of this paradox are reiterated in the PLA, where it is used to argue against the idea that mental concepts are determined privately:

> The impossibility of private language emerges as a corollary of his sceptical solution to his own paradox ... It turns out that the sceptical solution does not allow us to speak of a single individual, considered by himself and in isolation, as ever meaning anything (Kripke, 1982, pp.68-9)

The problem with this, as already stated in *PI* §202, is that if an individual is in charge of whether they have correctly applied a rule governing the use of a mental concept, then:

> ...whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'. *PI* §258

In other words, if I am judge and jury over whether a mental concept has been applied correctly, (i.e. if the rules are (in this sense) *privately* determined), then whatever 'seems right to me is right'; and this, for the reasons given above, is unacceptable.

Papineau records Wittgenstein's response to this as follows:

> Wittgenstein took it to be a condition on properly meaningful terms that their meaning be communicable. A coinage whose content must remain private to the coiner of the term is no coinage at all. (2011, p.181)

According to this line of thinking, Wittgenstein would be opposed to the very idea of phenomenal concepts: Phenomenal concepts are concepts which purport to refer to something that can only be accessed *privately* by the experiencer. However, as per the PLA, concepts must be *publically* accessible in order to count as meaningful. Therefore, there can be no phenomenal concepts. And as discussed, it is on this basis that Balog and Papineau respond with the Knowledge Argument: The PLA defends the view that concepts must be *publically* accessible. However, as per the Knowledge Argument, there are phenomenal concepts. Phenomenal concepts are *private*, in that they can only be accessed by the experiencer. Therefore, the PLA is mistaken.

The major flaw in Balog's and Papineau's argument is that the K-F interpretation is not actually endorsed by Wittgenstein. This point I take to be well-established after the long exegetical battles of the 1980s and 1990s (see e.g. Hacker, 1986, Stern, 1994, pp.424-432, Miller and Wright (edd.), 2002).[15] So I will not spare much space retreading well-worn refutations, except insofar as to briefly continue a thought developing from the criticism that Kripke fails to consider the second part of *PI* §201(see e.g. McDowell, 1984, McGinn, 1984, Pears, 1988, pp.467ff., for detailed statements of this objection). The subsection reads:

> It can be seen that there is a misunderstanding here from the mere fact that in the course of our argument we give one interpretation after another; as if each one contented us at least for a moment, until we thought of yet another standing behind it. What this shows is that there is a way of grasping a rule which is *not* an *interpretation*, but which is exhibited in what we call "obeying the rule" and "going against it" in actual cases.

In this crucial passage, Wittgenstein makes explicit that he does not accept the sceptical conclusion of the paradox. Rather, he rejects its key premise; he regards it is a 'misunderstanding' to think that rule-following requires interpretation. So Wittgenstein does not reject the place of rules within our practices, he does not concede to the paradox that 'no course of action could be determined by a rule'. Instead he stresses that we rethink the basis on which rules *are* followed. The K-F interpretation completely misses this rejoinder to the paradox.

This now begs the question, how *are* rules followed? The following is, I believe, an indicative remark of Wittgenstein's response to this question:

[15] Although see Kusch (2006) for a sophisticated defence.

> Does human agreement *decide* what is red? Is it decided by appeal to the majority? Were we taught to determine colour in *that* way? *Z* §431

The obvious answer to these questions is 'no'. It is all too convoluted a picture to fit with how we follow rules during the course of our busy lives. Of course, there is agreement (for the most part) about how to apply colour concepts to which kinds of objects, in which kinds of circumstances, and so forth. Otherwise the game would never get off the ground. However, agreement is not reached via an overt *decision* to do so; we don't take a vote and 'appeal to the majority.' Rather:

> "So you are saying that human beings decide what is true and what is false?"—It is what human beings *say* that is true and false; and they agree in the *language* they use. That is not agreement in opinions but in form of life. *PI* §241

> Language did not emerge from some kind of ratiocination *OC* §475

Agreement is reached out of 'form of life', not 'opinions'. Ignoring the extensive connotations of the term 'form of life', part of the message here is that rules are followed for the most part in virtue of something more akin to a *propensity* than to a ratiocination. In accordance with these remarks is the idea that rule-following is not the result of *interpreting* the rule in some cognitively prescient fashion. On most occasions, we just 'obey the rule *blindly*' (*PI* §219). Just as we are inclined to obey rules blindly, just as we are inclined to tacitly accept that others follow rules correctly, there is a similar ethos when it comes to our own use of concepts:

> I did not get my picture of the world by satisfying myself of its correctness: nor do I have it because I am satisfied of its correctness. No: it is the inherited background against which I distinguish between true and false OC §94, also see §§358-9

It is a consistent theme of Wittgenstein's to emphasise that it is our inculcation or 'training' (e.g. *PI* §§5-6, 198) from which we go on to act how we do and draw the kinds of distinctions we draw. The community rarely actively legislates over what counts as right or wrong concept application in practice. Agreement, insofar as it is required, is typically assumed to be sound unless there is good reason for suspicion otherwise. This is what distinguishes us from the fastidious shopkeeper of *PI* §1, who checks each word of the note in order to fulfil the request for 'five red apples'. There are certain things that we accept without feeling the need for much scrutiny. If rule-following is indeed like this, then it is not at all obvious that the existence of phenomenal concepts need be discounted by the K-F interpretation. Insofar as our use of concepts are publically checked, then, we might suppose that such checking procedures are

fairly lax whereby, say, all that is required to 'behave appropriately' and so pass the check is for us to avow possessing the requisite concept when questioned by a member of the community.

Balog and Papineau intend something significantly stronger than the lax public checks I refer to above. Indeed, they seem to require that the only sorts of checks capable of satiating public scrutiny are ones where the public have *direct* access to the concept. And possibly for good reason; it might be claimed that direct access is required in virtue of the nature of phenomenal concepts—their meaning is wholly attached to the *direct subjective* experience. In consequence, there is, effectively, no such thing as a phenomenal concept understood indirectly: the concept of an experience that is understood indirectly will be a different concept from the concept of that experience understood directly. In order to understand the phenomenal concept, one has to be the possessor of it. And, as per *PI* §§202 and 258, a concept cut off from the jurisdiction of the community in this way, can be no genuine concept at all.

I suggest that this desire for *direct* access is, at least from a Wittgensteinian point of view, misleading. It gets matters back to front. The important question is not whether the community can understand a concept without having direct access to it, but whether that lack of access is something that troubles the community. And the answer to this question, for Wittgenstein, is not predetermined by a priori general constraints on meaning but by the idiosyncrasies of the community itself. For Wittgenstein:

> Following according to the rule is FUNDAMENTAL to our language-game. It characterises what we call description (*RFM*, p.330)

It is the fact that we follow rules which counts as 'fundamental' to the game being played. It is not, pace the K-F interpretation, *the community* or *susceptibility to a public check* that is fundamental. That is, the nature of agreement over what counts as acceptable or unacceptable conditions for concept application are not predetermined according to some general constraint on meaning, such as susceptibility to a public check, but according to the form of life and whether or not it gives expression to such concerns.

Response 3:

Papineau and Balog's approach to Wittgenstein's notion of privacy can be stated bluntly: Wittgenstein rejects privacy. If Wittgenstein rejects privacy, and phenomenal concepts are private, it follows that Wittgenstein rejects phenomenal concepts. Therefore, Wittgenstein rejects phenomenal concepts.

The biggest problem for this approach is that it fails to recognise Wittgenstein's far more nuanced position in regards to privacy. Specifically, it conflates at the very least two distinct kinds of privacy: *logical* privacy and what I will term *non-logical* privacy. This distinction is drawn from the very opening of the PLA at *PI* §243, where Wittgenstein explicitly endorses the idea that there are some kinds of thoughts which are recognisably 'private'—we are given an example where one writes in a private code, another example where one talks to oneself in monologue, and even an example where there is an imagined society of people who only speak to themselves, in monologue. Of course, these cases differ: the private code is private because it is indecipherable, the internal monologue is private because others cannot hear it, and the society of monolinguists' thoughts are private, perhaps, because others cannot question what their thoughts are. However, all these examples represent perfectly meaningful uses of language where, in the requisite sense, such use is 'private', insofar as it is inaccessible to others. However, these examples are 'not what [Wittgenstein] mean[s]' when he speaks of a *logically private* language. As per *PI* §243: 'the individual words of this language are to refer to what can *only* be known to the person speaking; to his immediate private sensations. So another person *cannot* understand the language' (emphasis added). One reason why non-logical privacy differs from logical privacy is that although it is true that the thoughts expressed in the non-logical cases are in fact inaccessible to others, this need not be so. The private code writer could provide a key for his code, one could share one's thoughts publically, and the society of monolinguists could learn to speak in dialogue (perhaps with the help of the explorer). So although those cases of privacy *are not* accessible to others, it is *in principle* possible that they could be accessed. Part of what makes the logically private language interesting, then, is that it is supposed to represent an example where the meaning of that language *cannot* be understood by anyone but the speaker. It is only these logically private concepts which are the target of the PLA.

The root question, then, is whether Papineau is right to interpret Wittgenstein as committed to the idea that incommunicable sensation concepts are private in the logical sense, and thus to be rejected. To answer this, we can look to the short remark at *PI* §248:

> The proposition "Sensations are private" is comparable to: "One plays patience by oneself"

The rules of patience dictate that the game is played by one person, and in this sense, patience is a game that is played *privately*. There are, for instance, no two player versions of patience, (except insofar as one might derive some variant version where two or more people take it in turns to move). Clearly, however, patience is not a logically private game. Patience is a game with a common set of rules, rules that others can readily understand. I think this is the comparison Wittgenstein wishes to draw with sensations in *PI* §248: sensation concepts refer to feelings that can only be had by the individual who feels the pain, and so are private. However, similarly to patience, this privacy is *part of the game*: it is part of the (commonly understood) rules for 'having a pain' that they refer to sensations that are privately accessible only to the

person having the pain. This contrasts with logical privacy, where *the rules* are private too, insofar as others do not so much as understand the game one is playing. I return to the positive question of what logical privacy itself is in more detail in Part IV. The relevant point for now is that, for Wittgenstein, the privacy of sensations does not amount to logical privacy, insofar as those who do not have it are still capable of understanding it as a pain that the individual feels.

This brings us to the question: Is Marianna's 'Φ sensation' concept private in this relatively innocuous non-logical sense? In Papineau's defence, Marianna's 'Φ sensation' is not a straightforward instance of a sensation concept that follows rules which are readily understandable by others. Unlike the concept 'pain', Marianna is unable to communicate which sensation her concept refers to. So nobody, except Marianna, can know whether what she feels is a tingle, a pang, an ache, a throb, a tickle, and so on. They may even doubt whether Marianna is really referring to a sensation at all, given how she is incapable of communicating what her new concept refers to. Marianna herself might not even know what to call what just happened to her; she might, for instance, misname it as a burn, or a scratch. This all makes Marianna's concept increasingly inaccessible to others.

However, what Marianna shares with the early examples of *PI* §243 and the comparison with patience at *PI* §248, is that although she initially cannot communicate what her concept refers to, there is no barrier *in principle* to her being able to do so in the future. There are no a priori prohibitions on Marianna's coming to communicate in a publically understandable language what her 'Φ sensation' concept refers to, given propitious circumstances, i.e. circumstances where she can align her 'Φ sensation' with its counterpart in public language. All she has to do is ask her tormentors the right questions, or wait for the sensation to occur again, or observe someone else get a shock and then getting one oneself. There is even scope for saying that Marianna *can* communicate what her concept refers to in at least some publically understandable form—she knows that what she felt is sharp, painful, unpleasant, and so on. Perhaps, similarly to Hume's missing shade of blue, she might be able to interpolate that her 'Φ sensation' refers to tingling on those bases alone. So Marianna's concept is *in principle* intelligible to others, even if, as it so stands, it is incommunicable at the time at which she first forms the concept. In light of this, I think we should conclude that Marianna's privacy is to be classed as an example of non-logical privacy.

Bizarrely, in the midst of his exposition, Papineau comments that he personally supports a distinction between kinds of privacy. Furthermore, he thinks Marianna's privacy is a genuine case of non-logical privacy, privacy that is in principle perfectly intelligible. Papineau even references the argument presented by Hopkins (1974), which explicitly defends the view that Wittgenstein was well aware of such a distinction and only objected to 'in principle' private concepts. In the face of all of this, Papineau still concludes that Wittgenstein would object to Marianna's in principle publically intelligible concept (2011, pp.181-2). I must admit that the only reason I can find for Papineau drawing this conclusion, in face of all

this contrary evidence, is because of a dogged commitment to the bad arguments of the K-F interpretation.

Papineau goes on to levy a sophisticated charge against Wittgenstein, arguing that the PLA 'also places strong further requirements on the way such reference may be *secured* in the first place' (ibid. p.181, also see his 2002, pp.128-132). Papineau argues that Marianna's discovery that her Φ sensation is, in fact, what is commonly referred to as a pain sensation must be an empirical discovery; it is not a discovery that Marianna can make simply by reflection on the two concepts. But, Papineau argues, here's the rub; a posteriori identities cannot be achieved unless the two sides of the identity already have determinate meaning. If this is so, then 'Φ sensation' must have determinate meaning prior to its identification with the publically understood concept 'pain sensation.' And if this is right, then the point at which Marianna first coins her concept 'Φ sensation', her coining must have been established privately. And this, Papineau claims, is contrary to the standards imposed by the PLA (2011, p.182). Contrary to the above arguments, one cannot reply by arguing that her private concept is *in principle* intelligible to all, because prior to the discovery, Marianna had no way of telling whether her new concept was logically private or non-logically so. So far as Marianna is concerned before her discovery, it could equally have turned out that her 'Φ sensation' referred to a kind of sensation which was in principle incommunicable. In which case, she could not have set up the concept in the first place. And yet, even though for all Marianna knew her concept was logically private, it is still plausible to suggest that she forms a genuine concept.

This response takes us towards much deeper divisions in opinion about the nature of concepts, and about the 'where and when' semantics takes hold in the order of explanation concerning the origins of conceptual mastery. One question to pursue is to ask whether the discovery is straightforwardly empirical, whether Marianna's identification of her concept with the public concept arises simply as a result of her relating her private concept with the public one. As I suggest in my first main response to Papineau's interpretation, there is a question regarding the influence of deeper normative forces which underlie the ability for one to form a concept in the first place. And, if my suggestions there are correct, Marianna's acquisition of her 'Φ sensation' is, at least in part, already someway pre-prepared before she then goes on to make her discovery about the identity of it with red sensations. If this is so, Papineau's claim that Marianna's formation of her 'Φ sensation' concept is acquired entirely in private is belied by the far more complex normative situation in which Marianna finds herself. Marianna knows that there are things we call sensations, she has had many kinds of such feelings, is aware that sensations have distinctive qualitative feels, that vary in intensity, and so forth. Moreover, it seems plausible that Marianna is able to reflect on this and make rational inferences about how the kind of experience she is undergoing fits into these related aspects of experience. Arguably, then, Marianna's uptake of 'Φ sensation' and her consequent identification of it with 'red sensation' does not arise in some kind of semantic vacuum, but is

continuous with many other things Marianna knows about the world. Marianna's acquisition arises from within a normative framework of rules, not without.

Papineau sees matters differently. According to Papineau, normative rules '*follow* from the prior naturalistic constitution of content, and are not a precondition thereof' (2002, p.129). Applying to Marianna: Marianna is naturally endowed with concept-forming capacities which, given the appropriate circumstances in her environment, will cause her to form the requisite concept (as per a properly developed teleosemantic theory of meaning (see Millikan (1984), Papineau (1999)). Accordingly, she is able to form concepts even though she is 'not sensitive to *any* normative principles tying its use to public criteria' (Papineau, 2002, p.129, emphasis added). Under this theory, Marianna can, in a situation lacking any normative principles, set up a new concept by taking 'a potential experience concept' from an innate stock of potential concepts for types of experience, and 'lock[ing] it onto the type of experience at hand. Nothing more is needed' (Papineau, 2011, p.182).

If the arguments presented in my first objection are on target, then we should discount Marianna as representing a good example of someone who forms a concept in enough of a semantic vacuum to show that she is capable of acquiring her concept without '*any* normative principles' guiding her. In which case Papineau is left to answer the challenge: Can we find clear examples where concept acquisition *does* occur in isolation from normative rules? If such an example can be presented then we might have firmer grounds for rejecting the Wittgensteinian approach to explaining concept acquisition. I leave it until Part's III and IV to pursue this question in greater detail.

Response 4:

If the dialectic of argument is on track, then there is so far no persuasive reason to suspect that Wittgenstein would be opposed to phenomenal concepts. This still leaves us with the task of offering *positive* reasons for favouring the view that Wittgenstein would be at ease with the idea of phenomenal concepts. It is towards this question to which I now turn.

In an upcoming paper, William Child argues against the view that phenomenal concepts are a proper target of the PLA:[16]

[16] References to Child's paper are taken from a draft copy, set to be published soon.

> I shall argue for the unorthodox view that a Wittgensteinian approach to sensations and sensation language can accommodate a version of the idea that there are distinctive concepts of sensations that are available only to those who know what it is like to have those sensations (p.1)

To defend this unorthodox view,[17] Child cites examples suggesting that Wittgenstein expressed a fairly casual commitment to the idea that some concepts can be grasped without knowing what it's like to have them:

> I [never] have feelings of an invisible presence; other people do, and I can question them about their experiences (*PI, Part II,* §vii, p.184).

Under a natural reading of this example, we might be inclined to say, for example, Wittgenstein knows *that* such feelings occur when entering old houses, or *that* such feelings make one go cold and breathless. He knows *that* there is something it is like to have those feelings, even though he does not know what those feelings feel like. So Wittgenstein has more than enough semantic grounds to get by in a perfectly meaningful conversation with those who do have such feelings. With this point in mind, Child introduces the more interesting question in relation to Wittgenstein exegesis:

> ...would [Wittgenstein] accept that there are also concepts of experience that are available *only* to those who do know what it is like to have those experiences?' (Child, p.4, emphasis added).

This question revolves around whether Wittgenstein accommodates for the idea that there is a certain kind of knowledge (typically) available *only* through being appropriately acquainted with the right kind of experience. As discussed, critics argue that this scenario is surely in tension with the PLA because it suggests that securing the reference of a distinctive kind of concept is possible by relying squarely on acquaintance with the right kinds of experiences and the individual's innate cognitive powers to conceptualise the experience appropriately (see Papineau, 2011, p.182).

Child draws attention to another example:

> Remembering has no experiential content ... Would this situation be conceivable: for the first time in his life someone remembers something and says: 'Yes, now I know what

[17] The interpretation is 'unorthodox' in its attempt to accommodate for phenomenal concepts. Not only does this challenge Balog and Papineau, and therefore advocates of the K-F interpretation. It also opposes ideas defended by Hacker (2012), Hanfling (2001a), and Malcolm (1984b), who argue at length that Wittgenstein would have given short shrift to the idea that there is 'something it is like' to be conscious.

> "remembering" is, what it *feels like* to remember'. – How does he know that this feeling is 'remembering'? Compare: 'Yes, now I know what "tingling" is' (he has perhaps had an electric shock for the first time). (*PI, Part II*, §xiii p.196)

The main point of this remark is to criticise the idea that 'remembering' possesses distinct 'experiential content' (what I am calling 'phenomenal character'). What is notable for present purposes is Wittgenstein's contrast between 'remembering' and 'tingling'. 'Tingling' is here represented as an example of a mental (sensation) concept which *does* possess a distinctly phenomenal character, and so *is* tied up with distinctive qualitative properties, with *feeling a particular way*. So, again, Wittgenstein appears to be relaxed about the idea that some concepts get their meaning through referring to the phenomenal character of the experience; moreover, in this passage, he readily admits that knowledge of such concepts is based on acquaintance with the experience—the subject knows what 'tingling' is, only after having the shock for the first time.

Based on this ready acknowledgement of how sensation concepts possess distinctive phenomenal characters, Child suggests that the 'tingling' example can be naturally converted into a version of Knowledge Argument:

> [Let us assume] the subject already has a way of picking out the sensation of tingling: he can think of it, say, as the sensation one has when one has a mild electric shock. Then he has an electric shock for the first time; he experiences the tingling sensation for the first time. At that point he can truly say, or think, 'Now I know what tingling is; what it feels like to have an electric shock.' And, though Wittgenstein does not put it this way, it seems reasonable to say that the person is now in a position to think of the sensation of tingling in a new way; a way that draws on his own experience of tingling sensations (Child, p.5)

According to the scenario, the subject has *an* understanding of tingling sensations prior to the experience. He knows that tingling sensations are caused by electric shocks, that they are painful, and so on. Even so, in virtue of receiving the shock for the first time, the subject can now think of tingling in a new way, one that draws from the 'experiential content' of the sensation. This way of thinking was unavailable beforehand. If this scenario is plausibly conceived, it appears to show Wittgenstein at ease with the idea that there are concepts which can be acquired only after having had the relevant experience.

The tingling example so far stated may be criticised on similar grounds to the way in which the Mary example was criticised earlier. That is, one could argue that the scenario somehow fails to represent a satisfactory example for supporting a Knowledge Argument. Prefiguring this objection, Child introduces an extension to the 'tingling' case similar to the extension from Mary to Marianna (ibid. p.5). As with the

original scenario, we are to suppose that the subject has never felt a tingling sensation, but yet knows that electric shocks cause tingling sensations, that they are painful etc. However, in this new case, we give the subject a series of electric shocks, *but without telling him that this is what we are doing*. Child argues that in this case, similarly to Marianna, we can imagine the subject mastering a use for the concept 'Φ sensation', even though he doesn't realise that his 'Φ sensation' concept refers to the same sensation that other people refer to when speaking of 'tingling sensations'. He might be able to imaginatively recreate the Φ sensation, and make truth-evaluable judgements regarding it. Moreover, learning that his Φ sensations *are* tingling sensations seems like genuine epistemic progress, insofar as there is something illuminating about this discovery. It seems, then, that he can form a concept through '*only*' being acquainted with the experience, and this is possible quite aside from any ability to convey in publically understandable language just what his 'Φ sensation' concept refers to. He can seemingly do all of this without support from the community and in virtue of his natural powers of introspection. And crucially, if Child's line of reasoning is correct, this is all compatible with a Wittgensteinian rendering of phenomenal concepts.

I fully agree with Child's unorthodoxy here; there is clear textual support indicating ways in which Wittgenstein would be at ease with the idea of phenomenal concepts. I also think that if we allow for a broader understanding of phenomenal concepts, then we can continue this line of thought in new interesting directions.

Current literature on phenomenal concepts concentrates on what we might term '*basic sensorial experiences*', the majority of examples revolving around sensory stimulations as occur when e.g. seeing a red rose, hearing a new sound, or smelling a new smell. Papineau even goes so far as to explicitly defend a version of the idea that phenomenal concepts *are* a special variety of perceptual concepts (see his (2006, §3.2), and (2011, p.176);[18] there is, I think, more than a hint of classical empiricism to this position. This balance of emphasis on sensory experience derives, perhaps, from the influence of Jackson's (1982) paper, which explicitly takes issue with Nagel's (1974) definition of consciousness as 'what it is like to *be*' something. There, Jackson avers that philosophical focus ought to be rightly oriented over questions concerning the qualitative properties of conscious experience, rather than what it is like to *be* a conscious subject (see Jackson, 1982, p.130). I disagree with this limitation in philosophical focus, and believe it is a virtue of Nagel's paper that he latches on to what we might term as '*non-sensorial experience*,' experiences that do not obviously refer to the specifically sensory aspect of the experience.

[18] In his (2006, footnote 12), Papineau alludes to the possibility of which I am about to focus heavily on, namely that there exist nonperceptual (or at least *non-sensorially perceptual)* but yet phenomenal concepts. In my view, accepting this premise allows us to envisage a far wider scope for the role that phenomenal concepts play in our cognitions.

I want to use the term '*perspective*' to incorporate this wider sense by which to understand experience, as contrasted perhaps with the prior focus on 'perception'.[19] By perspective, one includes sensorial experiences, but also describes the non-sensorial features of experience, which at a local level might include the beliefs and attitudes as expressed by an individual, and on a more global level as the *world-view*, or *form of life* in virtue of which one assesses, and takes part in, their environment. At present, there is not room to defend this position with any depth, however the basic intuition is as follows: Consciousness is not merely the result of successions of qualitative perceptions; rather, these successions are organised and arranged within a perspective which situates them, assigns them relative significance, and responds to them in ways it deems appropriate. This perspective is formed, in part, as a result of upbringing within a particular complex normative framework, or socio-cultural milieu.[20] Radically divergent perspectives give rise to failures in understanding and communication even though the sensory stimulations might be of equal or similar viscosity. In at least some cases, then, understanding requires sharing a perspective, not just a perception.

It is my view that Wittgenstein sometimes latches onto this sort of idea. In the following passage, Wittgenstein certainly seems committed to the idea that there is a certain kind of understanding available only to those who share in a form of life:

> We also say of some people that they are transparent to us. It is, however, important as regards this observation that one human being can be a complete enigma to another. We learn this when we come to a strange country with entirely strange traditions; and, what is more, even given a mastery of the country's language. We do not *understand* the people. (And not because of not knowing what they are saying to themselves.) We cannot find our feet with them. (*PI*, *Part II,* §xi, p.190)

Here we might imagine that the traveller has acute anthropological understanding; he understands *that* these exotic people believe such and such, display such and such a ritual, revere these idols, and so on. And yet, for all of this, we can imagine that the traveller might still feel that he doesn't understand what it is like to be them, that he doesn't quite grasp their way of life.

To develop this thought with a biographical example; on a number of occasions Wittgenstein confessed his inability to believe the Christian doctrines that his Catholic friends, Elizabeth Anscombe and Yorick Smythies, believed (see e.g. Malcolm, 1984a, p.60 and Monk, 1990, pp.463-4). Specifically, he stated that he could not believe in the resurrection, transubstantiation in the Eucharist, or the idea that Jesus is Lord. However, Wittgenstein claimed that this inability was not due to the fact that he 'believed in the opposite'

[19] Nagel (1974) focuses towards phenomenal character in this wider sense, although he equivocates somewhat.
[20] This is not to mention the more obviously physical-causal conditions required for forming a perspective.

to them over such matters (*LCA*, p.53, also pp.54-9). Rather, his complaint concerned his inability to utter expressions of these beliefs with any genuine meaning; e.g. the word 'Lord', he says, could convey meaning for him 'only if [he] lived *completely* differently' (*CV*, p.33, see ibid. pp.32-3). Just what Wittgenstein doesn't understand about his friends religious beliefs is an important question that I cannot here pursue with requisite depth (although see Schroeder, 2007, and Tripodi, 2013, for interesting discussion). However, what is clear is that Wittgenstein is claiming that there is a sense in which he does not *fully* understand the meaning behind his friends' uses of the word 'Lord,' and that this inability was due to the fact that, as one commentator puts it, there are some religious 'truths whose accessibility conditions include certain requirements as to the attitude of the subject' (Cottingham 2010, p.223).

These examples show Wittgenstein expressing commitment to the idea that there is a kind of understanding available *only* through sharing the attitudes and beliefs, the 'form of life', of those who do embrace such things. Furthermore, this understanding can be represented as a distinct kind of knowing 'what it is like', one that is unavailable to the uninitiated. Without sharing in the form of life, whatever wealth of descriptive information and understanding one might accumulate about, for example, transubstantiation, one will still be in the dark when it comes to all the peculiarities which positively contribute to an understanding of what it is like to experience in those kinds of ways. We might say that without being able to share in these experiences, one can *sympathise* with those who do experience them, but never truly *empathise*. Another way of expressing this is to say that there are some aspects of our experiences that are *private*, in the sense that they are beyond the understanding of those who do not sufficiently share in our form of life. Such differences in perspective give reason to consider a new dimension to Wittgenstein's notion of privacy, and the ways in which we might have access, and fail to have access, to the inner lives of others.

PART III: Externalist interpretations of the PLA

If Wittgenstein is no opponent of phenomenal concepts, the following question repeats itself: What *does* count as a genuine instance of logical privacy, as per the PLA? In this Part, the externalist interpretations of Wittgenstein's response to this question will be critically discussed, as found in Pears (1988, 2008) and Child (2011, forthcoming), and elements of which are found in Hacker (e.g. 1993a, 1993b).

Externalist interpretations of the PLA:

According to Child's interpretation of Wittgenstein:

> What individuates sensations? What makes a sensation the kind of sensation it is? When we reflect on that question, Wittgenstein thinks, we find it natural to think that sensations are individuated by their subjective, introspectible character. And we find it natural to think that its subjective character is a purely intrinsic feature of a sensation: a feature whose identity is entirely independent of anything to do with the subject's behaviour or external circumstances. ... Such a view of sensations has been dominant in the history of philosophy; and it is popular in contemporary philosophy. And, Wittgenstein thinks, it exerts a natural appeal on anyone who reflects on the nature of sensations. (2011, pp.151-2)

For Child, the point of the PLA is to oppose the idea that it is the 'intrinsic' feel of the sensation that matters, picked out 'purely' in virtue of its 'subjective, introspectible character'. It is this view which commits those under its influence to (knowingly or not) advocate the existence of a *logically* private language.

The question of *who* holds this view of sensations is undoubtedly a large one. Child and Pears, (and Hacker in similar discussions, see his 1993a, pp.1-2, 8-12, 17-20), indicate some historical sources: Descartes is mentioned, presumably insofar as he supposed that he could successfully detach his thinking entirely from the external world and still retain meaningful thoughts. Similar views are represented by traditional empiricism, such as Locke and Hume, and the more recent sense-data theories as found in e.g. Russell, Carnap, and Ayer. It is expressed by Frege and Kant (see Hacker, ibid. pp.17-19). The specific theories which articulate such views are perhaps less significant than the 'natural' way of thinking about sensations out of which such theories arise. The leading idea seems to be that unlike concepts which refer to physical objects, and whose meaning derives from their attachment to objects in the 'external' world, our sensation concepts are meaningful in virtue of referring to objects in our 'inner' mental world, i.e. to internal sense-impressions. One upshot of this view is accessibility; whereas it can be readily challenged whether the bird I am looking at is a crane or a heron, it seems that the only authority on what I am feeling, or the way things appear to me, is myself. Only I have direct access to my thoughts and sensations, and so only I can know the contents of my mind; the best others can do is to surmise or infer about those contents. It is not difficult to see how this 'natural' conception of mental content leads toward more profound philosophical puzzlements: If only the individual has access to the contents of their own mind, then perhaps everybody experiences colours differently (*PI* §272), perhaps nobody except oneself even has sensations at all (*PI* §420). Such considerations draw us towards a conception of mental concepts which specifies them as fundamentally *private*: If mental concepts refer to inner thoughts and sensations, and if those thoughts and sensations can only be accessed by the individual who has them, then it appears to follow that the meaning of such concepts can therefore only be accessed by and

known to the individual who has them. If this is so, then it seems that we are left with the rudiments of a *private* language.

Child's interpretation of Wittgenstein's response to this view is worth quoting at length:

> The target of the discussion of private language in *Philosophical Investigations* and elsewhere is a view of sensations on which the identity conditions are completely independent of any links to external circumstances or behaviour ... Against this view, he argues that there is no way of individuating the kind of experience one has when something looks red except in a way that ties the character of the experience to the objective circumstance of seeing something red. [One might have an experience of something as red when there's no red thing there (on occasion)]. But, Wittgenstein insists, that does not commit us to the idea that experiences have phenomenal characters whose identity conditions are entirely independent of all links to external circumstances, physiology, or behaviour (forthcoming, p.6)

Child is unequivocal: for Wittgenstein, the point of the PLA is to show that the identity conditions of sensations are necessarily tied up with 'external circumstances, physiology, or behaviour'. (Going forward, this idea is what is broadly referred to as *externalism*). As outlined in my fourth response in Part II, Child successfully argues for why Wittgenstein would think there is nothing inherently problematic about concepts which are not capable of being communicated to others, as in the case of Marianna. Communicability is philosophically-speaking inconsequential to whether or not Marianna exercises a concept. Rather, what is essential is whether Marianna's concept succeeds in referring to some relevant feature of the external world.

Similar views are echoed in Pears' recent monograph:

> [Wittgenstein's] leading idea is that the language in which we report sensations owes its meaning to their connections with the physical world and cannot survive separation from it. (2006, p.41)

Elsewhere:

> ...reports of sense-impressions are expressed in a language that is essentially dependent on reports of their causes in the physical world, including the bodies of the speakers. (ibid. p.50)[21]

[21] Pears references the source for this claim as found in Wittgenstein's remark in *Philosophical Remarks*, §47*:*

And:

> ...the function of sense-impressions is to give us information about the physical world in which we have to live our lives, and so the meanings of our reports of our sense-impressions are preserved by their success in performing this role. (ibid. p.51)

The target of the PLA, under these interpretations, is the view that the identity conditions of sensations are understandable independently of the external causes by which they are ordinarily associated. Colours are properties of objects, and colour concepts are meaningful in virtue of referring to them. Feelings of pain are typically caused in response to external stimuli, and pain concepts refer to parts of the body. The PLA does not deny that one can, on occasion, identify the sensation without the appropriate external correlate (e.g. phantom limbs), or have the external correlate without identifying the sensation (e.g. localised anaesthetic). Rather, the target is the idea that sensations can be understood as '*entirely* independent' of any links to the external environment. It is concepts which aim to refer to sensations under this guise which represent the real example of logically private concepts, and the kind to which Wittgenstein's PLA is fundamentally opposed.

What is wrong with the disputed view, according to the externalist interpretation? Pears argues that if one restricts appeal to only sense-impressions and memory of sense-impressions, then one takes away the independent standard necessary for establishing a reliable basis for meaning (see Pears, 2008, pp.43-4, 50-5):

> a colour-word like 'blue' could not be given a stable meaning by a would-be private linguist, whose only resource would be his visual impressions of blue without any regular connections with blue physical objects. He would have to rely on the remembered similarity of a sequence of sense impressions ... Wittgenstein's criticism is that this would not be a reliable basis for linguistic regularity [i.e. meaning, see ibid. Chapter 2]. The only way to get a reliable basis would be to use physical objects independently known to be blue. (ibid. p.44)

> For since language only derives the way in which it means from its meaning, from the world, no language is conceivable which does not represent the world.

As a side-note, I am cautious that this remark should be taken as an authoritative voice of Wittgenstein's later philosophy. The notes from *Philosophical Remarks* are sometimes a mixture of *Tractatus*-type views with his newer approach to philosophy. The above quote, as it stands, has a disputably-sounding *Tractarian* ring to it. Even a remark such as at *PI* §23, where Wittgenstein lists ways of using words that *do not* aim at representing the world, might be seen to be in tension with it. But this point belongs to other important issues.

One reason this is problematic is that without reference to the physical world it becomes unclear how one could distinguish those impressions which are genuine from those impressions which only appear to be. After all, both genuine and false impressions present themselves in the same way, i.e. as impressions. Without the external world, 'there would be no criterion of correctness' (ibid. pp.44) and so we lose the ability to make this crucial distinction.

Pears considers Ayer's response to this argument (Ayer, 1986, pp.75-80). In short, Ayer argues that the necessary reference to independent external cues can be preserved, it is just that our understanding of those external cues is to be reformulated in ultimately sensory terms. This is achieved by checking a current sense impression against a group of sense impressions. A false impression would simply be one that fails to stand up against the tribunal of the group. In this way, Ayer argues, the distinction between 'genuine' and 'false' impressions can be preserved, it is just that this distinction is to be drawn from within the sense-impression framework.

Pears rejects this reformulation as an 'inadequate substitute', arguing that it represents an 'essential change' to the original requirement (op. cit, p.53). In support, he references Wittgenstein:

> Are the rules of the private language *impressions* of rules?—The balance on which impressions are weighed is not the *impression* of a balance. *PI* §259

> "Well I *believe* that this is the sensation S again."—Perhaps you *believe* that you believe it!
> *PI* §260

For Pears, Wittgenstein's point seems to be that no amount of checking against impressions can make up for the severance with the physical world which the sense-data theorist posits. The point here is not to defend the sceptical idea that 'we can never know' whether our memories are genuine, but to reject the idea that memory-images *alone* could supply such a genuine basis for meaning. Referencing another remark of Wittgenstein's:

> For example, I don't know if I have remembered the time of departure of a train right and to check it I call to mind how a page of the time-table looked. Isn't it the same here?—No; for this process has got to produce a memory which is actually correct. If the mental image of the time-table could not itself be tested for correctness, how could it confirm the correctness of the first memory? ... Looking up a table in the imagination is no more looking up a table than the image of the result of an imagined experiment is the result of an experiment.
> *PI* §265

As these quotations point towards, sense-impressions, belief-impressions, or memory-impressions cannot guarantee that one remembers correctly, for no matter how convinced one may be that the impression conveys the right information, unless that conviction is backed up by what is in fact the case (i.e. the external world), there is always open the possibility that one only appears to be seeing, believing or remembering correctly.

There is scope for responding to Pears' criticism here: According to Pears' Wittgenstein, impressions alone are not a sufficiently reliable guide for establishing meaning. Checking sense-impressions against sense-impressions can never get to the truth, no matter how vast the tribunal of impression-based evidence one might appeal to. What is required is the confirmation of those impressions by reference to physical objects. In response, I would argue that Pears' argument shows only that we are to admit that our recollections are fallible, and that we have no neat refutation against the spectre of full-blown memory/ impression/ belief scepticism. It does not show that one could not set up a language on the basis of impressions alone. Checking the train time-table that springs to one's imagination is not conclusive evidence that one knows the right time. This is to be admitted. However, it may be that such images are the best one can hope for in the situation. Of course, it is no substitute for looking at the actual time-table—looking at actual time-tables offer much firmer grounds for believing that a train will depart at the stated time. It is not immediately clear to me why a sense-data theory cannot maintain this distinction, just so long as we allow that 'looking at the actual time-table' is understood as occurring via a mental image, and does not of itself provide a cast-iron guarantee that the time-table is actual. In which case, such images are the best one can hope for in *any* situation. So we might argue in favour of Ayer, that his theory can maintain this distinction between having a memory and thinking (falsely) that one has a memory. And he can maintain that it is sometimes the case that one mistakes the two, and even makes those mistakes many times without realising it. Indeed, on this view, it is a coherent (if far-fetched) thought that all one's memories are false. So long as one accepts these provisos about one's fallibilism, then I see no problem endorsing the idea that we rely on impressions. Moreover, if memory were to be so utterly distrusted, it is difficult to see grounds for *ever* placing trust in our memories. Overall, reliance on memory/sense/belief-impressions does not entail that the distinction between genuine and false memory collapses or that memory cannot be checked.

Accepting this response, one further problem is bounced back for Pears' Wittgenstein. Whereas the Ayer-type view may be accused of failing to countenance the ordinary ways in which we make contact with the world, Pears' Wittgenstein fails to countenance the ways in which we often struggle to make those connections, and ways we might altogether fail. It might be argued against Pears' Wittgenstein that he fails to pay due heed to this fallibility. If sensation concepts are formed as a matter of reference to the physical world, then how to explain those cases where we are convinced that our concepts have been formed in such a fashion only to find out (or perhaps never finding out, as per the sceptic) that they only appear to

have done so. Ayer's problem seems to be distinguishing real sensations in a world of supposed appearances, Pears' problem, conversely, is being able to distinguish appearances in a world of supposed physical objects. One might wonder whether these are two sides of the same problematic coin (and I suggest as much in Part IV). If there are definitive problems with Ayer's view, I think we should look for them elsewhere.

Child presents a different argument. To this end, he quotes *PI* §258:

> Let us imagine the following case. I want to keep a diary about the recurrence of a certain sensation. To this end I associate it with the sign "S" and write this sign in a calendar for every day on which I have the sensation.—I first want to remark that a definition of the sign cannot be formulated.—But still I can give myself a kind of ostensive definition.—How? Can I point to the sensation? Not in the ordinary sense. But I speak, or write the sign down, and at the same time I concentrate my attention on the sensation—and so, as it were, point to it inwardly.—But what is this ceremony for? For that is all it seems to be! A definition surely serves to establish the meaning of a sign.—Well, that is done precisely by the concentrating of my attention; for in this way I impress on myself the connexion between the sign and the sensation.—But "I impress it on myself" can only mean: this process brings it about that I remember the connexion *right* in the future. But in the present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'.

The language-game of sensation 'S' in §§258ff. discusses an attempt to establish the meaning of a new sensation word 'S' in a scenario where there is (i) no natural behaviour through which to express having S and (ii) no descriptive vocabulary through which to describe what S is, or what S is like. For example, it can't be shown where the sensation occurs, by e.g. rubbing your arm or crying out etc., because this would be a part of natural expression, and so violate condition (i). And it can't be said whether S is pleasant or painful, nor say how long it lasts for etc., because these are descriptive terms in a readily understandable public language, and so violate condition (ii). Wittgenstein regards S as an instance of logical privacy, and so too regards it as ultimately incoherent.

According to Child, *PI* §258 has Wittgenstein opposing the idea that a sensation concept can be formed through the minimal conditions of having a sensation and consequent naming of it. Under this reading, sensation S is used by Wittgenstein as an ex-hypothesis example where we are to imagine the subject as drumming up a meaningful concept through the powers of introspection alone, by establishing correct use through 'concentrating attention' on a particular sensation and proposing to use the word 'S' to refer to all future instances of that type. The precise question of the section, as Child asks it, is: '*what* is it for

something to be the same kind of sensation as the one that was originally called 'S'?' (op. cit p.154). Here, we cannot just assume that there are preset standards by which we can establish what is to count as the same. Indeed, because 'S' is meant to be private, the individual cannot appeal to resources that are available in ordinary, public, language (ibid. p.156):

> it would not help [for the private linguist] to say that it need not be a *sensation*; that when he writes 'S', he has *something* — and that is all that can be said. But 'has' and 'something' also belong to our common language (*PI* §261)[22]

Rather, the private linguist is required to 'conjure up absolutely everything he needs for defining his private words entirely from his own, introspective resources' (ibid. p.157). And the idea that this is possible, Child claims, is for Wittgenstein 'pure fantasy' (ibid. p.157).

But what is it about this situation that makes it so fantastical? Child thinks we should contrast what we can say about private colour concepts with what we can say about private sensation concepts. Similarly to the conclusions discussed in Part II, the self-formed colour concept is in principle communicable to anyone, given propitious circumstances. That is, it is possible to imagine that someone could set up meaningful practices for use of a colour-word without knowing how to communicate about it publically (ibid. p.157). What makes it possible to form such a concept, according to Child, is that colour concepts refer to a 'relatively enduring property' (ibid. p.158), so the individual will be able to form genuine distinctions between correct and incorrect applications, between the *appearance* of the colour on an occasion and the actual presence of the colour. The same stability, he argues, is not true of sensations. By contrast, Child claims that sensations are 'simply too ephemeral' (ibid. p.158) to afford us the possibility of formulating concepts about them in isolation from reference to external circumstances.

Child argues that Wittgenstein provides this external reference in *PI* §270, with his example of the manometer. *PI* §270 is a development of the example of *PI* §258, where the subject is imagined to have learnt that whenever he has S there is an increase in his blood pressure. So the feeling of S is identified with rising blood pressure. This identification, Wittgenstein comments, 'is a useful result'. For Child this means we have our required external cue: 'what an S-type sensation *is* is the feeling of my blood pressure's rising' (Child, forthcoming, p.7). So, the next time the subject feels S again, he can comment that his blood pressure is rising. This is not to claim that sensations *are* behaviour, that the sensation *is* blood pressure rising. For Child, it really is meaningful to say that it is the *feeling* that matters:

[22] Please note that I return to the topic of whether it really is the private linguist who is the correct subject of paraphrase in objection (f) of Part IV below.

> Having an electric shock is not merely a physical and physiological phenomenon; a normally-functioning person who is given a mild electric shock really does *feel* something; she has a sensation. (p.7)

The point is restricted to the twin idea that a sensation cannot be conceived as something with an intrinsic, subjective, purely introspectible character, and that reference to the external environment is a constitutive part of what it is to experience the sensation. This result, if we are to believe Child's interpretation, is what saves the sensation from ephemerality, and affords it a stable basis upon which to conceptualise what S is all about. The next time the subject feels S again, he can reliably comment that his blood pressure is rising.

Expressivist interpretations of the PLA:

There is a question regarding what consequences we might expect from acknowledging a necessary tie between sensation concepts and external circumstances, physiology, or behaviour. As argued, externalists believe that sensation concepts are necessarily tied to external circumstances, physiology, or behaviour. Under this view, it is natural to suppose that the character of the tie should in some way reflect the dynamic between the behavioural expression and the concept which aims to characterise it. *Expressivism* is a natural corollary of externalism in this respect. Broadly speaking, expressivism defends the idea that meaningful expressions of sensation concepts—or, *avowals*—are such in virtue of the concepts relation with the behavioural *expression* of the felt sensation. Expressivism is set in contrast to 'descriptivism', which defines that relation in terms of the expressed concept being the result of some introspective process of identifying the sensation and reporting on it. Both Pears and Child defend expressivist interpretations of Wittgenstein's positive view of the relation between sensations and sensation concepts (Child, 2011, pp.166-175, Pears, 2006, pp.55-7). Quoting Wittgenstein:

> Here is one possibility: words are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behaviour.
> "So you are saying that the word 'pain' really means crying?—on the contrary: the verbal expression of pain replaces crying and does not describe it. (*PI* §244)

> Being sure that someone is in pain, doubting whether he is, and so on, are so many natural, instinctive, kinds of behaviour towards other human beings, and our language is merely an

> auxiliary to, and further extension of, this relation. Our language-game is an extension of primitive behaviour. (*Z* §545)

Expressivist interpretations are represented strongly in the literature, from a variety of perspectives (see Ginet (1968), Fogelin (1987), Hacker (e.g. 1993a, 1993b), Wright (1998), and Bar-On (2004)). It would be mistaken to think these commentators defend one uniform set of views on the matter; far from it. Even so, I think we may distinguish two broad versions: *strong* expressivism and *nuanced* expressivism. According to the strong expressivist interpretation, avowals are better considered as the *embodiment* of the behaviour, rather than a report of it. In this manner, behaviour is *part* of the concept. So whilst we use words in such situations, we could just as well have groaned, grimaced, or pulled facial expressions, for this is ultimately the same thing. As Fogelin puts it:

> Crying is not a report about our feelings of pain, but an expression of them … the same can be said for the remark "I have a pain in my arm". Saying this is also part of our pain behaviour, not a comment upon it (1987, p.170).

What makes this interpretation *strong*, is that avowals are considered as so tightly tied to natural behavioural expression, that they are entirely dependent upon them. In order for there to be an avowal, there must be behavioural expression (though, of course, this isn't the case vice versa). Fogelin pays only lip-service to their conceptual separability (ibid. p.240, footnote 1), but emphasises that:

> Wittgenstein speaks of "one possibility" [§244] for the explanation of sensation talk, but nowhere offers any other, and the things he says next [§§245-6, §§256ff.] presuppose that something *very like* this account must be correct (ibid. pp.169-70).

And Fogelin certainly proceeds by assuming that Wittgenstein only really ever had this one possibility in mind. In a similarly *strong* interpretation, Bar-On argues that, for Wittgenstein, first-person avowals are 'in no way thought to represent or be epistemically grounded in a subject's *judgement* or *belief* about some state of affairs' (Bar-On, 2004, p.230). That is, they do not serve to describe one's present condition in the first place. Thus:

> [I]t is misguided to regard the special status of avowals as a consequence of their recognitional access that subjects have to their own states of mind, and there is no need to seek any secure epistemic basis on which avowals are made. (ibid. p.231)

For Bar-On's Wittgenstein, avowals are not based on a process of recognition and report, but expressions of the sensation itself. As expressions, avowals are inherently *open* to view instead of hidden away to be

inferred of or surmised about; they *display* their meaning, not provide a report on it. As such, it is 'as inappropriate, conceptually speaking, to assess their epistemic credentials as it would be to assess a moan or cry or a laugh in terms of its evidence, correctness, or reasons' (ibid. p.231).

There are problems with such strong views. As Bar-On argues, this form of expressivism struggles to account for the semantics of avowals as they are employed in conditionals and other indirect senses, (see Bar-On's extension of the Frege-Geach objection to expressivism (ibid. p.233-40)). Moreover, expressivism has a problem in accounting for avowals relating to complex inner experiences. For example, suppose I avow that I am currently feeling mildly smug that my opponent resigned in our game of chess the previous Thursday. It is difficult to see how this feeling can clearly translate into a type of behavioural expression. And even if we can, there is a question regarding why we would want to, when the simpler explanation is that I am simply describing how I feel. Furthermore, not all sensations are nearly as intense as feeling pain. Rather, it seems that some avowals are capable of being reported; one can coolly *relay* that they are feeling relaxed in a bath, and *describe* that they are experiencing pleasurable feelings without having to translate those avowals into an instance of natural behaviour expression. That is, it seems to be part of the 'depth grammar' (*PI* §664) that on at least some occasions avowals are reports, not expressions. If this is acceptable, then avowals and behavioural expression are not tied up in the way the strong expressivist maintains.

There is, however, a weaker interpretation of Wittgenstein's supposed expressivism, which avoids many of the problems faced by the strong expressivist. This *nuanced* expressivism is explicitly defended by both Child and Hacker. The nuance is not to claim that all avowals are only expressions of behaviour. Instead, avowals are understood as *grounded* or *'rooted in'* natural expressive behaviour:

> The vast majority of our desires have no natural, prelinguistic behavioural expression; but their expression is nevertheless rooted in the primitive behaviour of striving to get or crying for something or other ... Once the primitive linguistic extension is grafted onto the natural expressive behaviour, further linguistic extensions grow. For the mastery of a language opens up the possibility of ever more subtle, refined, and linguistically differentiated pain-*behaviour* (Hacker, 1993b p.92).

Nuanced expressivism is weaker than strong expressivism insofar as it does not tie avowals directly to current expressive behaviours. Rather, the tie may be recognised only via some very indirect link back to natural expressions. As long as somewhere in the history of the use of an avowal there is some contact with natural expressive behaviour, then there is no issue with the avowal being many steps removed from it. Importantly however, the inclusion of *linguistic* extensions enables those original expressive behaviours to develop in a vast number of directions. Linguistic extensions not only allow us to divorce our use of

sensation concepts from the direct expression of sensation behaviours, but allow us to generate wholly new experiences altogether; experiences that are unavailable to those without the requisite linguistic capabilities. As Child exemplifies:

> the budding wine connoisseur's acquisition of a sophisticated vocabulary for describing the tastes of different wines goes hand in hand with her coming to experience wines in richer, more complex ways. It is not that she learns to describe more accurately the experiences that she already has; rather she comes to have different experiences (2011, p.173)

Nuanced expressivism can readily accommodate for these semantic features of avowals—it can explain their use in conditional statements, their ability to be coolly reported on, the complex games we learn how to play with them etc. There is no requirement to try and force a translation of avowals into natural expressive behaviour.

This weaker tie undoubtedly better reflects the sophisticated ways in which expressive behaviour relates to use. Nuanced expressivism twinned with externalism results in a detailed set of views attributed to Wittgenstein's opinions on the philosophy of mind and psychology. It is one that keeps an eye open to both the organic nature and 'primitive' origins of language and meaning (*Z* §545), but also to the normative structures that arise out of complex language-using creatures in specific socio-cultural contexts.

Objections to externalism and nuanced expressivism:

Whilst I fully agree with Child's unorthodox view that Wittgenstein is no opponent of phenomenal concepts, I disagree with his positive proposal for how Wittgenstein *would* characterise them. In the rest of Part III, I provide reasons to suspect that the externalist interpretation is not in itself a convincing argument. In Part IV, I suggest that externalism is not representative of Wittgenstein's actual thinking on such matters, and offer an alternative proposal.

Objection 1:

If sensation concepts are meaningful in virtue of their ties to external circumstances, physiology, or behaviour, then a sensation without such ties is, by this definition, not meaningful. And this is indeed the conclusion that the externalist draws from the PLA. One example which might run counter to this line of thinking is generalised anxiety. Diagnosed as a medical condition or 'disorder', generalised anxiety may be

regarded as a feeling, but a feeling that is not directed towards any particular object or circumstance, and seems to arise 'for no apparent reason';[23] it is, as its name suggests, a 'generalised' feeling. In this way it is unlike e.g. 'tingling', which is caused by something externally determinate, i.e. an electric shock. It is not clear what external object or circumstance to which we can point to for generalised anxiety, or indeed whether there is such a cause. And yet, I will assume that generalised anxiety is a real phenomenon.

One response might be to deny that there are no underlying external cues which bring the feeling about. There may be physiological causes due to hormone imbalances, or more macroscopic causes such as unrealised work stresses or lifestyle dissatisfaction. This response is of course highly controversial, and there is scant room to go into it in appropriate depth here. I will only suggest that I am not convinced that *all* cases of generalised anxiety can be so easily reduced in these ways. Another response might be to reject the idea that generalised anxiety has *no* external cause, instead suggesting that it has *many*. In other words, the generality of the anxiety is such that it is to be recognised as a result of a great number of anxieties relating to various external causes formed under one banner. I think this is a more promising line, although a credible answer will perhaps only be fully realisable through the results of careful psychological studies. I am no psychologist and so will only comment, with reservation, that it is well-known that attempts are often made to tackle external causes supposedly contributing to the anxious mental state, only for the individual to return to that state without those causes. In which case, we might think that those supposed factors are perhaps only superficially related to the underlying sense of anxiety—they are a consequence of the generalised anxiety as opposed to its cause. An alternative response might be that the generalised feeling of anxiety may yet be 'rooted in' naturally expressive behaviours, even if it is not an immediate expression of them. I think this story can be told, and with some plausibility. For example, we can readily imagine some primitive sources for feelings of anxiety, arising perhaps as an evolutionary response to life on the plains of the Serengeti, in order to avoid dangers of potential predators. Or we might interpret a baby's crying as a natural expression of anxiety at his parent's departure. These primitive bases may serve as the original mould from which more sophisticated forms of anxiety arise, such as those well expressed by modern thinkers, Durkheim's *anomie* for instance, or perhaps Kierkegaardian *angst*. This may be right. Even so, I think it is a fair question to ask why we should expect that the connection with some long-distant conceptual ancestry has much to do with explaining generalised anxiety as experienced in its more sophisticated form. The link might very well be there, but it doesn't explain very much. If the reason for bringing it in rests only in the desire to avoid objections, then this is *ad hoc* and therefore a bad reason. I am not suggesting that this is the reason. I am unsure what the reason might be. As such, I move on in the hope that other objections lend support to the claim that we need not appeal to conceptual ancestry in order to explain the semantics of phenomena such as generalised anxiety.

[23] For reference, as of June 2014, see http://www.nhs.uk/conditions/anxiety/Pages/Introduction.aspx

Objection 2:

My second objection centres on Child's claim that private sensations are 'simply too ephemeral' to act as reliable sources from which we might construct sensation concepts. As discussed earlier, Child contrasts sensations with colour experiences (see ibid. pp.157-9). Colour experiences, it is argued, *do* offer a sufficient basis from which to construct colour concepts. This is supposedly because colour experiences, unlike sensations, possess a 'relatively enduring property'. This durability enables them to be compared with other objects of the same colour, it allows one to order objects inasmuch as they accurately match that colour, and so on (ibid. p.158). Private sensations, Child argues, are quite different:

> ...a private sensation is available for inspection only as long as he can hold it in his attention; he cannot then retrieve the very same sensation and consider its qualities again. Nor can he compare the features of one private sensation with those of another that is enjoyed at a different time ... [T]here is nothing to sustain the practices of sorting and classification that would be needed if the private linguist were to be able to establish standards of correctness for the use of his private sensation words (ibid. pp.158-9)

In short, it is not possible to 'retrieve the very same sensation' or compare it with new candidate private sensations. The sensation is too ephemeral for that.

An important point of contention is *why* private sensations must be regarded as ephemeral. I do not see why they need to be. If the sensation is capable of having an impact on the subject, enough so that its occurrence is registered in the first place, then why should the subject be incapable of recognising it when it, or a similar such experience, appears again? One might think of how this is possible with non-private sensations. Many non-private sensations intuitively qualify as lively enough to remember after having been experienced just once; for example, eating an extra hot chilli, having a toothache, or perhaps the early stages of falling in love. These sensations, though in various ways fleeting, are lively and durable enough such that if they have chance to arise again, one will be able to think e.g. "Ah, I've felt *that* before". That is, one will be more than capable of recognising the new sensation as the same or similar in kind to the one experienced the first time around. Such sensations are anything but ephemeral.

So what should make *private* sensations any different? Why can't they be remembered? The natural answer for Child has to be that non-private sensations have ties to external circumstances which provide the basis upon which they are remembered. E.g., the sensation of heat is tied to the chillies, the toothache is

tied to the tooth, and one's love is tied to one's lover. S, we are to imagine, has nothing external by which to remember it by and therefore slips away every time we might wish to capture it.

One question to pursue is to ask whether the idea of experiencing an acceptably non-private sensation without its external counterpart makes it any less conceivable. It seems to me that we should not immediately expect this to be the case. After all, the sensation surely still *feels the same*: the heat still feels hot, the toothache still painful. It is just that now such feelings arise and dissolve without interventions from external influences. If this is imaginable, then I fail to see what is wrong with the thought that sensations might have ever only arose in that way. And if we can imagine that, then we are imagining a meaningful concept without external manifestation; something which the externalist suggests that we cannot do.

A similar criticism was put forward in an objection to behaviourism by Putnam in the early nineteen-sixties. (So far as I can see, the objection applies equally well to behaviourism as it does to externalism). Putnam sought to defend the idea that:

> From the statement that 'X has a pain' by itself *no* behavioural statement follows—not even a behavioural statement with a 'normally' or a 'probably' in it (1968, p.9).

To this end, he asks us to imagine a community, he calls them the 'Super-Spartans', where the inhabitants feel pain sensations but do not behaviourally express them or report on them in any way. One feels pain, one may even give a secret name to it, but one does not, and will not ever, express it to anyone else. (They might even wonder "Who knows, perhaps all the other Super-Spartans have the same thing", but of course they would never express this). Putnam's point is that, for all its peculiarity, there is nothing contradictory or incoherent about this scenario. Again, the example represents how we can imagine meaningful concept acquisition and application without external manifestation.

One response to these objections might be that imagining a non-private sensation without its usual external ties is different from trying to imagine a sensation that has been in total isolation from them from the very beginning. When imagining pain without external ties, we (as philosophers conceiving the scenario), can make sense of it. It is just feeling a pain without external manifestation. Borrowing from *PI* §257: 'what is presupposed is the existence of the grammar of the word "pain"; it shews the post where the new word is stationed'. In other words, we can use the ordinarily-understood grammar of the concept 'pain' and apply it to a case where the grounds on which the concept is formed are altered. This is why it is not difficult to imagine. However, this option is not available to the logically private sensation. The private sensation has no such ordinary grammatical 'post' to begin with and so cannot be stationed elsewhere. This is why S cannot be imagined. S fundamentally differs from unexpressible pain sensations

due to its lack of an original station. However, this station is not determined as a result of a lack of external ties. If what I have argued above is correct, it is still the case that one can strip away all the externals and retain the feeling. We can imagine the occurrence of a sensation independently from its external circumstance, and we can imagine forming a phenomenal concept in response even though there is nothing external to support those cognitive activities. If this is so, then it is mistaken to think that unexpressible pain sensations are incoherent. As the non-private cases demonstrate, a central component of constructing a sensation concept is *how the sensation feels intrinsically*. These issues are further discussed below and in Part IV.

Objection 3:

Let us suppose that the previous two objections miss their target. Here is another: As discussed earlier, Child claims that 'tingling' is a word tied up with external circumstances, specifically with having an electric shock. And this thought is maintained even in the wake of a scenario where it is not possible to communicate what one means to others—all that is required is the right external contact. We might object, however, that even if external circumstances are necessary, they do little to *explain* the phenomenal character of phenomenal concepts. The concern is that even if there must be external ties, the ties have little explanatory force in the face of the stark fact that the sensation *feels a (distinctive) kind of way*. If the externalities cannot perform this explanatory function, then a full explanation of those concepts must lie somewhere else. And for this we need to include reference to the *intrinsic feel* of the sensation.

The bizarre thought-experiments involving zombies and wholesale colour spectrum inversions have been variously designed to provide objection to the idea that external circumstances can tell us all there is to tell about the content of sensations and experiences. The uniting idea is that: for all that external circumstances *can* tell us about sensations and experiences, they cannot tell us everything. There is always the possibility that *something* is different, that something diverges, from what we ordinarily take it to be. Chalmers (1996) and Kirk (1999) discuss zombies, imagining a scenario where nobody except for oneself possesses phenomenal character, even though others display all relevant behaviours to suspect that they do. Shoemaker (1982) and Block (2007) explicitly use the example of widespread spectrum inversion to object to Wittgenstein's philosophy. In Block's argument in favour of widespread colour inversion, he provides an elaborate thought-experiment. We are asked to imagine the following: a person grows up with normal colour vision. He then has an operation which alters his colour vision, things that previously appeared red now appear green and vice versa. The person then goes about his life, and learns to apply the terms red and green as he had done before the operation; he never calls the grass 'red', or the sunset 'green', and so forth. Over time, he comes to use the words 'red' and 'green' as familiarly as he had done

before, even though he is, in fact, seeing invertedly. Finally, the poor person suffers amnesia and so forgets his earlier experience, including the operation. At this stage, the individual uses ordinary language in just the way everybody else does, his behaviour is indistinguishable from those who are not colour inverted, and he has no way of telling that he is inverted. He has no reason (so far as he knows) to suspect that he has ever been inverted, and so on. Nevertheless, his spectrum *is*, in fact, inverted. He still sees red where everybody else sees green and vice versa (ibid. pp.91-95).

In defending a Wittgenstein argument against these objections, Child writes the following:

> What individuates sensations? What makes a sensation the kind of sensation it is? When we reflect on that question, Wittgenstein thinks, we find it natural to think that sensations are individuated by their subjective, introspectible character. And we find it natural to think that its subjective character is a purely intrinsic feature of a sensation: a feature whose identity is entirely independent of anything to do with the subject's behaviour or external circumstances. So, we think, it is perfectly possible for two people to be subject to all the same external stimuli, and to be exactly alike in every behavioural respect, but for the subjective character of their sensations to be entirely different: it is possible, for example, 'that one section of mankind [has] one visual impression of red, and another section another' (PI §272) ... Such a view of sensations has been dominant in the history of philosophy; and it is popular in contemporary philosophy. And, Wittgenstein thinks, it exerts a natural appeal on anyone who reflects on the nature of sensations; though, as we shall see, he thinks the appeal of this view depends on misunderstanding the 'grammar' of our sensation language. (Child, op. cit. pp.151-152)

According to Child's Wittgenstein, the idea that wholesale inverted spectra are possible is based on a misunderstanding of the true 'grammar' of sensations. As we have seen, this 'grammar' is argued to necessarily involve extension to external stimuli. The 'natural' way to think is misguided.

Other Wittgensteinians defend Wittgenstein against the examples in different ways. Horwich argues that if the possibility of wholesale spectrum inversion is raised 'relative to normal conditions, and with respect to ordinary language characterisations of experience, the answer is, trivially, that there is no such possibility' (2012, p.204). If, however, the question is raised relative to the philosopher's supposed 'private experiences' then the question is revealed to be a pseudo-question because 'there are no such private qualities and private terms' (ibid. p.204). Non-normal conditions do not count because these are conditions philosophers mistakenly try to impose on the basis of a faulty assumption about the privacy of the inner. Kilverstein presents Wittgenstein's worry as being that:

> ...if spectrum inversion were widespread, the terms we use in reporting how we experience colour would, as he puts it, 'lose their use' ... If spectrum inversion were widespread, there would be no agreement about how to use colour terms in describing one's experience ... Without any agreement on which usages count as normal, talk of how things look with respect to colour will be meaningless (2008, p.37)

For Kilverstein's Wittgenstein, if spectrum inversion were a live possibility, then reports on experience would 'lose their use' because nobody would agree about what how to use colour terms correctly and so such talk would become 'meaningless'. Tanney argues against the idea that there could be a physical, functional and behavioural duplicate (a zombie) that somehow lacked consciousness. She argues that once we provide a detailed behavioural explanation of the supposed zombie, it is difficult to make sense of just what the zombie is supposedly lacking (2004, pp.176-8). Shoemaker appeals to Wittgenstein's remark at *PI* §272:

> The essential thing about private experience is really not that each person possesses his own exemplar, but that nobody knows whether other people also have *this* or something else. The assumption would thus be possible—though unverifiable—that one section of mankind had one sensation of red and another section another.

Shoemaker claims this remark is intended to represent a '*reductio ad absurdum* of the notion of 'private experience' ... that it implies that this 'assumption' might be true' (1982, p.328). See also Child's take on *PI* §272 in the previous quotation, and Horwich (op cit., p.204 ft.30) for similar interpretations. All agree that Wittgenstein is here talking *ex hypothesis*.

I think this is wrong. Unless one is already persuaded by the externalist interpretation, there is less reason to regard Wittgenstein as fundamentally (in principle) opposed to zombies or widespread spectrum inversion. Moreover, I think we can appeal to textual remarks to defend the view that Wittgenstein was not necessarily opposed to the possibility of widespread spectrum inversion or zombies. As I read *PI* §272, Wittgenstein is here genuinely conceding that it is possible (though unverifiable) 'for one section of mankind' to experience one colour and another section another colour. The 'essential thing' is not whether 'each person possesses their own exemplar', it is whether people are concerned about it. Kilverstein's worry depends on the idea that widespread inversion would have to have practical implications on practices, such that if there were such inversion, there would be no agreement. It seems to me, however, that there could very well be widespread inversion in co-existence with agreement in the use of colour terms. It would just be (unbeknownst to the participants) that the agreement was based on false premises. The practice need not break down simply in virtue of the possibility. What matters (again, the 'essential thing') is whether the individuals who co-operate in those practices allow the possibility of

inversion to overrun them. But the answer to this is not determined on a priori grounds, but according to whether the 'forms of life' that use them are concerned enough about it (e.g. *PI* §217). Certainly in *our* practices, the thought is relatively inane; it *might* be the case, but "so what?" I would argue that the possibility is, for Wittgenstein, a live one, but, frankly speaking, it doesn't matter enough to let it disrupt our lives. If it did, the practice of colour identification probably wouldn't last long.

If Wittgenstein is theoretically indifferent to the possibility of widespread colour inversion, then this suggests that we can see it as having a role *within* ordinary language (albeit a fairly exotic branch thereof). If this is so, we have grounds for rejecting Horwich's presentation of the matters, which depends on the idea that such scepticism is not part of our 'normal conditions'. If by 'normal' Horwich thinks we should be restricted to considering only mundane everyday thoughts and conversations (or something of the sort), then I think he attributes to Wittgenstein too strict a conception of what counts as 'normal'. Many thoughts and ideas that arise in ordinary (non-philosophical) life stray beyond the mundane, and it would be Orwellian to suggest that meaningful language must be confined within those reaches. A zebra wearing a Mexican hat surfing on an ironing board is not a 'normal' thought, but is well within the bounds of ordinary language and thought; perhaps spectrum inversions and philosophical zombies are no less bizarre. As will be defended in Part IV, these kinds of thoughts are of a different category to what we are asked to imagine when imagining sensation S; S is not just bizarre, it is plain incoherent. And this is what makes it a genuine instance of a logically private language. However, more on this later.

Consider the following remark relating to the possibility of zombies:

> But can't I imagine that the people around me are automata, lack consciousness, even though they behave in the same way as usual? – If I imagine it now – alone in my room – I see people with fixed looks (as in a trance) going about their business – the idea is perhaps a little uncanny. But just try to keep hold of this idea in your ordinary intercourse with others, in the street, say! Say to yourself, for example, "The children over there are mere automata; all their liveliness is mere automatism." And you will either find these words becoming quite meaningless; or you will produce in yourself some kind of uncanny feeling, or something of this sort. (*PI*, §420)

What can be noted is how readily Wittgenstein concedes to the possibility of zombies in the first two sentences. If he regarded the idea as incoherent, he would be likely to have pointed it out. But he doesn't; rather he opposes it on the grounds that it is simply something that we cannot maintain with seriousness outside of philosophical reflection on those matters. In this sense we might read the passage as Humean in its concession to the logic of sceptical thought, as measured against the fact that our nature compels us to believe regardless. Importantly, Wittgenstein's response is not to answer the problem of other minds

by *justifying* our beliefs. Neither is it to suggest that the consideration is somehow incoherent. As with the response to colour inversion, the main flaw in the example is its inability to purchase a grip on us; it is not that we cannot make sense of it. In this manner, I think we might read Wittgenstein as allowing that such examples have *a* place in our practices, even if only for a relatively circumscribed time, with little genuine force.

There are further remarks which display a similar attitude in response to privacy and accessibility issues. We can see this at play in the discussion between *PI* §§273-278. For instance, Wittgenstein asks the question:

> What am I to say about the word "red"?—that it means something 'confronting us all' and that everyone should really have another word, besides this one, to mean his *own* sensation of red? (*PI* §273)

Note that he never answers this question. And in *PI* §§277-8, Wittgenstein admits that there is a sense to which our concepts refer to private colour sensations, known only to ourselves:

> But how is it even possible for us to be tempted to think that we use a word to *mean* at one time the colour known to everyone—and at another the 'visual impression' which *I* am getting *now*? ... When I mean the colour impression that (as I should like to say) belongs to me alone I immerse myself in the colour. (*PI* §277)

> "I know how the colour green looks to *me*"—surely that makes sense!—Certainly: what use of the proposition are you thinking of? (*PI* §278)

As *PI* §278 remarks, if a coherent story can be told regarding the use of propositions about private colours, then Wittgenstein is open to hearing it. My suggestion is that accounts of zombies or Block's example of spectrum inversion *do* represent just those kinds of stories.

One final remark, *Z* §536:

> I may know that he is in pain, but I never know the exact degree of his pain. So here is something that he knows and that his expression of pain cannot tell me. Something purely private.

In the parenthesis to this remark, Wittgenstein does criticise the idea that there *has to be* reference to 'degrees of pain' within our concept of pain. Some uses make no reference to degree. I think this is right.

We often don't use the term with such a precise measure in mind. On the other hand, we can equally imagine how this is sometimes, indeed potentially *often*, the case. It can be admitted to be fairly ordinary to suppose that our knowledge of others pains is only an approximation, that we sometimes get it wrong, or that we sometimes fail to realise the intensity of others feelings. Again, such considerations lead us to suppose that externalities will not take us 'all the way' in terms of understanding the internal states of others.

Overall, the appellation of Wittgenstein as an externalist forces commentators to oppose spectrum inversion and zombie examples. In my view, the examples are coherent, and they do represent a challenge to externalist theories. It would be a mistake to oppose the idea that the most important thing about sensations precisely *is* their 'purely intrinsic feature', that is, *how they feel* to the subject. This feature is integral to the grammar of the concept of sensation. As a consequence, the inverted spectrum *is* possible, zombies *are* conceivable. If Wittgenstein objects to these consequences, then I believe he is in big trouble. However, I don't see why we should regard Wittgenstein as in principle opposed to zombies or inverted spectra, or to other less far-fetched examples where externally undetectable differences are less dramatically conceived. A Wittgensteinian can allow that zombies and inverted spectra give an insight into the sort of access and inaccessibility we have to others, and our epistemic boundaries, whilst also recognising that these insights are the thin edge of a wedge, when it comes to understanding our mental lives more widely.

To summarise the desired aims of Part III: If the argument is so far on track, then Wittgenstein should not attempt to tie sensation concepts to external circumstances, physiology, or behaviour in the manner suggested by the externalist interpretation. There are ready examples where the extrinsic can be separated from the intrinsic without loss of meaning. None of this is to contradict that Wittgenstein is correct to emphasise that there is *a* tie. In fact, it is entirely fair to say that Wittgenstein was a master of descriptive psychology, at discussing the intricate relations between behaviour and our use of concepts. However, if his philosophy of psychology extends to the claim that there *must* be such a tie, then I believe he is mistaken for the reasons given above. An essential aspect of our understanding of sensations and experiences is how they feel *intrinsically*.

PART IV: Logical privacy, non-logical privacy and the real target of the PLA

In this section, I *finally* move on to my attempt to positively answer the questions about what counts as logical privacy, what counts as non-logical privacy and therefore about what the overall point of the PLA

really is. Contrary to widespread interpretation, I argue that Wittgenstein intended the logical privacy of sensation S to be *contrasted* with the privacy of sensations without external counterpart. Under my reading, unextended pain sensations do *not* represent instances of logical privacy. I also attempt to provide a positive take on what non-logical privacy amounts to, when we include the idea that external circumstances don't explain everything. I suggest that recognising the sense in which privacy *is* a part of our everyday experience represents an interesting topic for philosophical investigation.

Logical privacy:

According to my reading of the PLA, the first genuine example of a truly logical private language is the introduction of sensation S at *PI* §258. As such, and in opposition to the externalist interpretation, unexpressible pain sensations are *not* examples of logical privacy. The chief difference, which I shall go on to discuss, is that whereas unexpressible pains are coherent, the point about purportedly *logically* private sensations is that they are, ultimately, *incoherent*. This shift in emphasis leads to a different treatment of *who* Wittgenstein regards as mistaken in their attempt to characterise S as a sensation, and *why* they fail.

Arguing for an externalist or expressivist interpretation relies on accepting the premise that Wittgenstein's early remarks about pain in the PLA (e.g. *PI* §§244-6, 256-7) are to be considered *comparatively* to his remarks about the logically private language, specifically with sensation S (*PI* §§243, 258ff.). Wittgenstein is regarded as saying that the idea of there being a concept 'pain' without expressive behaviour is similar to what he says about sensation S. Hacker, for instance, comments upon the conclusions drawn in *PI* §261, a remark exclusively about sensation S, by explicit comparison to *PI* §257, a remark about pain:

> The private ostensive definition of the word 'S' in the private language cannot be identified as a definition of a sensation-word by invoking the grammar of 'sensation' in the public language to determine the grammatical post at which 'S' is to stand (cf. §257). (Hacker, 1993b, p.5)

The comparison is evident at the end of the quote: '(cf. §257)'. Hacker takes Wittgenstein to be arguing against the idea that the words 'pain' and 'S' could mean anything in a scenario where there is no way to ever express in what circumstances one has S or pain. Both 'S' and 'pain' are equivalently meaningless insofar as saying or thinking that you have S or have pain by using a word 'S' or 'pain' solely by virtue of naming (ostensively defining) the supposed sensation in isolation from any outward display of those sensations, is senseless. Hacker claims that Wittgenstein's indulgence of the language-game where there is pain but no pain expression is to be understood as something considered only *ex hypothesis* possible; it is

indulged in, in order to show how upon investigation it doesn't make sense (ibid. p.59). One cannot imagine the word 'pain' being meaningful in a world where there is no pain expression or where there is no external circumstance to which one can tie the sensation to. Therefore, for Hacker, the example of pain in §257 is similar to when Wittgenstein indulges in the possibility of a logically private language such as sensation S even though he really believes this to be impossible also. Both the example of unexpressible pain concepts and the example of sensation S are to be understood comparatively, and both are regarded as examples of *logically* private concepts.

My main exegetical contention is that the externalist interpretation draws the wrong conclusions about the relation between what is said in *PI* §§244-6 and §§256-7 about exhibiting pain behaviour, and what is said about sensation 'S' in *PI* §§258ff. In short, I argue that §§244-6 and 256-7 are more illuminatingly considered *in contrast to* §§258ff. §257 focuses on what it would be like to be in a world where the relation between the concept *pain* and the outward expression of it is broken. §258 focuses on an example where one tries to name a sensation that is logically incommunicable. If I am correct, these are two very different investigations.

Under my reading, Wittgenstein is asking us to *contrast* sensation S with what has so far been said about the imaginability of pain sensations. Sensation S is a purported 'sensation' that bears *no connection* to any sensation concept we do understand. Given this starting premise, if we try to talk of sensation S as a sensation, if we try to justify it even as a sensation instead of, say, an emotion or a thought, we either (a) begin introducing ways in which the sensation is no longer an instance of logical privacy—insofar as it is now bearing connections with other concepts we do understand. Or (b), our introduction of it *as a sensation* becomes incoherent—as Wittgenstein says, to say so much even as 'that one 'has' it is' (§261) brings it away from the logically private domain, so presumably one can't even say of it this much. The result is that we can't conceive calling something a sensation that bears *no relation* to what we really in fact recognise as a sensation. This example contrasts with pain that is unexpressible or untied to external circumstance insofar as that sensation is still understood as feeling a certain kind of way, and so bears those connections with other concepts that S lacks. Sensation S doesn't succeed in getting this far in our imaginations. And it is this that separates acceptable (and understandable) privacy, from genuine logical privacy. The externalist's mistake is to conflate these two kinds.

I offer six reasons to justify my *contrastive* interpretation, hoping that each reason somewhat lends support to the others:

(a) It is noteworthy that pain is explicitly discussed from §244-257, but is not further discussed until much later (with the one brief exception of §263). §§258-271 concern 'S', and then Wittgenstein moves onto a discussion about colour experience for a number of sections up to §280. Only then

does the example of pain return at §281, and even then only in a context where he is discussing behaviourism (§§281-88). So why ought we to suppose that the remarks about pain are similar, in the above mentioned manner, to the remarks about S, especially as Wittgenstein nowhere himself draws the parallel?

(b) §257 itself may be read as a response to the interlocutors attempted understanding of Wittgenstein's position, which is stated at the very beginning of the section:

> [Interlocutor's voice attempting to emulate Wittgenstein's:] "What would it be like if human beings shewed no outward signs of pain (did not groan, grimace, etc.)? Then it would be impossible to teach a child the use for the word 'tooth-ache'".

The rest of the section seems clearly written in the form of an attempt for Wittgenstein to redress the interlocutor's misappropriation of his own views; and notice that the interlocutor's remarks are distinctly externalist. Wittgenstein might well be read as pre-empting and rejecting such a position in this very section.

Hacker muddies the division between what the interlocutor is saying and what Wittgenstein is saying in §257 (op cit. pp.59-60). As mentioned above, he treats Wittgenstein's response to the interlocutor as being that of getting him to consider more foundational difficulties than just being unable to teach the child the meaning of 'tooth-ache'. He also puts in the interlocutor's mouth the worry that the speaker in the example couldn't communicate his thoughts to others (ibid. p.59). But only the above quoted remark is in "…" marks, indicating this to be the voice of the interlocutor, whereas the rest of the section is without them, and so one assumes that the consequent voice is that of Wittgenstein redressing the interlocutor, not the interlocutor himself. Throughout his exegesis of §257, Hacker repeatedly switches between the voice of the interlocutor and that of Wittgenstein in a way that does not tally with how it occurs in the text.

(c) More attention should be paid to the fact that Wittgenstein nowhere outright denies that there is an experience of pain, nor that the child (being a genius) couldn't invent a word for it, nor even that there isn't a practice (given that there is a teacher), but only that it isn't *by fact* outwardly expressed or expressible. He merely questions the point, or 'purpose', of naming ones pain in a world where you are never going to express it or communicate it to anyone. And he is right, what would be the point? But he nowhere states that you couldn't *in principle* do it; he nowhere denies that the word you use to name the pain is still potentially understandable as meaning pain all the same. In short, he nowhere says this idea *lacks sense*. Sensation S on the other hand *does* lack sense. It fails to even get so far as to have a point or purpose. The failure to successfully imagine S is

not due to a lack of one's imaginative powers, but due to a lack of a coherently-conceived idea being put forward in the first place.

(d) Consider the last part of §257:

> When one says "He gave a name to his sensation" one forgets that a great deal of stage-setting in the language is presupposed if the mere act of naming is to make sense. And when we speak of someone's having given a name to pain, what is presupposed is the existence of the grammar of the word "pain"; it shews the post where the new word is stationed.

These are difficult remarks to break down. Hacker takes them to indicate that, for Wittgenstein, naming something necessarily requires a background or 'stage-setting' in order for the naming ceremony to meaningfully take place. As outlined above, he has in mind the stage-setting of behavioural expressions of pain; without these, the naming just doesn't make sense.

I do not think this interpretation takes proper account of the second sentence of the quoted remark, i.e.: *when we speak of someone's having given a name to pain, what is presupposed is the existence of the grammar of the word 'pain'*. As discussed in my second objection to externalism in Part III, this seems to indicate that Wittgenstein is clarifying what we do when we, qua philosophers looking at an imagined scenario, try to think of doing something like naming a pain (hence the 'when *we* speak of' aspect of the remark). When we, qua philosopher, think of this situation we immediately put into it our understanding of the grammar of the word 'pain' (something we all clearly understand). So the situation of §257 is perfectly imaginable; §257 represents only a case where there is no expression of pain amongst a community but where someone in that community somehow does this strange naming of pain and doesn't tell anyone about it. Wittgenstein is saying that when we think of this, we immediately put forward some understanding of the situation. As Wittgenstein says: '[the grammar of the word 'pain'] shews the post where the new word is stationed'.

We ought to *contrast* this situation with how we think about S. Trying to name an indefinable sensation, a sensation that occurs in isolation from anything we understand and any concept we possess, is meaningless, not because we can't express it to others, but because we, qua philosophers, don't even understand it ourselves. In short, S, unlike pain, precisely has no such 'post'. This is what makes it very different and a real instance of a case where 'one doesn't know what one should imagine here'.

(e) The immediately succeeding sections after §258 are remarks applicable only to sensation S, not the sensation of pain considered in §257. §261, for example, rejects the idea that the word 'S' can stand for a sensation, this is because the words 'sensation' and 'has' are words from the common stock of publically understandable language. One cannot say that S is like any other sensation, or that S feels a certain way, or even that it lasts for a certain period of time etc. because this stops it from being logically private in the requisite sense of §243. The point of sensation S is that it cannot be defined, and is logically incommunicable. This therefore includes defining it in terms such as 'it is similar to a pain sensation, but less sharp, and also it feels somehow more pleasant', for this puts it squarely back in a publically understood language. Wittgenstein even emphasises in the same section that one would just like to let out an inarticulate cry, but even this is something that occurs in an understandable language. Sensation S is unlike anything we've ever experienced. Contrast this with the pain example of §257. We know perfectly well what pain feels like, what it would be like for someone to feel pain but not express it, and so on. S cannot get nearly so far in our thoughts.

(f) As mentioned in footnote 23 of Part III, Child's suggested subject of paraphrase in *PI* §261 is that it is the private linguist who is incapable of imagining sensation S (2011, pp.156-7). The reason for the failure, as per Child, is that 'sensations are too ephemeral' and our memory and powers of introspection are not up to the task of remembering accurately. If this were so, the limitations are placed on the linguist himself; we might imagine that, were his introspective powers better, he may have been capable of forming a concept of the sensation.

There is an alternative answer as regards to *who* is making the mistake about S here. In short, it is not the private linguist who is mistaken. Instead, I suggest the following:

> it would not help [for us, *qua philosopher attempting to envisage the scenario*,] to say that it need not be a *sensation*; that when he writes 'S', he has *something* — and that is all that can be said. But 'has' and 'something' also belong to our common language. (*PI* §261)

According to this view, by 'us, *qua philosopher attempting to envisage the scenario*', Wittgenstein is placing the limitations on our *qua philosophers* attempts to conceive the scenario successfully in the first place. And it is we who are being scrutinised. This turns the object of the investigation around; it is not an attempt to provide constraints and rules for how we (in ordinary life) use concepts but an attempt to criticise the way we *qua philosophers* attempt to explain those uses. The investigation is thus focused not on the concepts themselves, but on philosopher's misappropriations of them.

In this way, we can regard the PLA not as a proof negating the possibility of a private language. Rather, it is as Stern (whose views bear some semblance with my own) puts it: 'Wittgenstein's principal aim is to get us to see that we only seem to understand the notion of a private language: that it falls apart on closer examination. In other words, we don't understand the notion of a private language well enough to be able to clearly state its nature, and then argue from such a premise to a conclusion concerning the impossibility of such a language' (2010, p.50, also see his 2004, p.174).

The opposition, then, is to the idea that we can conceive of S in the first place. It isn't as if *were his introspective abilities any better* the private linguist would be able to form a credible sensation concept with S. Rather, it is mistaken to think that there is a genuine naming of a *sensation* going on in the first place. A sensation is a concept that is meaningful in virtue of its place within a publically understood language, and when one tries to strip away all those elements of it that make it meaningful in the first place, one cannot retain the idea of a meaningful concept. This is a general constrain on what it is for a concept to have meaning: I cannot claim to have bought a 'car', but when asked to describe it, I say that it is a car with no wheels, not made of metal, has no seats or steering wheel, it cannot be driven, it has no engine, no make or model etc. In this instance, when I strip away all those elements which make up my concept 'car' and still aim to retain there being some sense to the notion, it is ultimately I who am mistaken in determining it *as a car*. A car is such precisely because it is something *with* those connotations in public language and cannot be deemed as such without them. Likewise, to say that there can be a sensation without any of the regular connotations by which to make it understood as a sensation is a misguided enterprise.

If these objections are on point, they veer towards a particular answer regarding the target of the PLA. I argue that one of the chief conclusions of the PLA, specifically the remarks about sensation S, is the idea that many of the concepts that philosophers suppose to be established through the having of an experience and consequent naming of it are in fact the product of a far more complicated set of relations than assumed. A scenario truly devoid of anything barring the intrinsically felt experience or sensation is not so simple as we might at first admit. We inevitably impute into our considerations concepts that are not so privately established (as in the private ascription of colour or pain). And when we do try to imagine such logically private cases, cases without them, we are unable to meaningfully state how it is what we want to say it is in the first place. The real point of the PLA is to stress ('remind') that our understanding of concepts is had in virtue of their possessing a certain substantive content that constitutes what we mean by them. This point is emphasised by the failure of sensation S, which attempts to act as a sensation

concept stripped of all such content through which it might be understood. However, when such attempts at application lack content, it soon follows that judgements based on them are not just contentless, but also meaningless. This reasoning can be upheld without resorting to the idea that concepts must be tied to external circumstances.

Metaphilosophy and methodology:

To move on: In asking whether Wittgenstein should be read as committed to externalism, it is important to look outside the confines of the PLA. Considering the remarks:

> And we may not advance any kind of theory. There must not be anything hypothetical in our considerations. We must do away with all *explanation*, and description alone must take its place. (*PI* §109)
>
> Philosophy simply puts everything before us, and neither explains nor deduces anything.—Since everything lies open to view there is nothing to explain. (*PI* §126)
>
> If one tried to advance *theses* in philosophy, it would never be possible to debate them, because everyone would agree to them. (*PI* §128)
>
> When philosophers use a word—"knowledge", "being", "object", "I", "proposition", "name"—and try to grasp the *essence* of the thing, one must always ask oneself: is the word ever actually used in this way in the language which is its original home?—
> What *we* do is bring words back from their metaphysical to their everyday use. (*PI* §116)

Remarks such as the above suggest something like the following: Wittgenstein's ambitions were not to produce new philosophical theories expounding the real essence of knowledge, or the mind, or personal identity and so forth, but to dissolve confusions which have led philosophers to assume that finding such 'essences' is a worthy philosophical ideal to aim for in the first place (*PI* §118-9). Such suppositions arise, Wittgenstein claims, from fundamentally misunderstanding how language works, and the therapeutic task Wittgenstein sets himself is to 'bring words back' to their everyday use, by reminding philosophers about the multifarious uses that constitute our applications of concepts and language, and from whence their original meaning is derived. Insofar as Wittgenstein has anything positive to say, it is trivial reminders about common usage which nobody (save for philosophers 'captured' by their misguided ideals (*PI* §115)) would deny.

What do these remarks have to do with the discussion about sensation concepts? Wittgenstein is advising that we recognise the multifarious ways in which such concepts can be applied, without trying to subsume them under one overarching theoretical framework. What is essential is that *we* as philosophers reflect (describe) the ways in which these concepts are used in our lives. To limit the use of sensation concepts to external circumstances or public checkability is almost inevitably going to be guilty of just the sort of exercise that Wittgenstein is condemning in these passages.

Many commentators take a dim view of these metaphilosophical remarks. For instance, Kripke:

> Had Wittgenstein—contrary to his notorious and cryptic maxim of §128—stated the outcome of his conclusions in the form of definite theses, it would have been very difficult to avoid formulating his doctrines in a form that consists in apparent sceptical denials of our ordinary assertions (1982, p.69)

Kripke clearly thinks we can separate Wittgenstein's good philosophy, which is a sceptical denial of ordinary assertions, from his bad, from his 'notorious and cryptic' maxims.[24]

Few Wittgensteinians seem prepared to stare straight into the mirror and say that Wittgenstein put forward *no* philosophical theses (see Fogelin, 2009, chapter 1, for good discussion). I am not entirely sure I can stare into that mirror myself.[25] However, I do think the metaphilosophical remarks provide the necessary backdrop for how we should be thinking about other aspects of Wittgenstein's philosophy. They are not to be excised from an interpretation of Wittgenstein's core 'doctrines' about rule-following and meaning. They are a functioning part of the machine.

As discussed in Part III, Child draws externalist conclusions from the argument of the PLA. How does this fit with the metaphilosophy? On the face of it, they are in conflict; externalism, however you look at it, is a theory. It is not a trivial reminder, but a highly controversial philosophical theory of a quite traditional kind. As Child admits, the 'natural appeal' (2011, p.152) is towards the idea of concepts being meaningful according to their intrinsic character, not their exterior. In a discussion concerning Wittgenstein's metaphilosophical views and his philosophy of mathematics, Child thinks interpreters should bite the bullet and admit that Wittgenstein's overall position is 'genuinely conflicting' (2011, p.134). The implication here is that this applies *pari passu* to his philosophy of psychology.

[24] I assume *PI* §128 equally refers to surrounding remarks, as found in e.g. *PI* §§89-133.
[25] The most convincing attempt to do so that I have come across is Kuusela (2008).

I find this conclusion hard to believe. Biographically, that *Philosophical Investigations* was constructed over sixteen or so years, that it was the product of intense thought and energy, not to mention intensive revision and editing. Moreover Wittgenstein's own passionate attachment to his metaphilosophical remarks belies the attribution of such glaring rudimentary inconsistency to his work. Even if the textual evidence were slim, much better, in my opinion, if we look for that evidence which gives the author the benefit of the doubt.

In my view, Wittgenstein's main concern was not to constrain concepts, but to criticise philosophers' misappropriations of them. The aim of providing a surveyable representation is, to some extent, heuristic; it is a means to dispelling philosopher's excessive generalisations of explanations that apply well under particular conditions. The mistake Wittgenstein attributes to philosophers, philosophers that would include externalists, is their concern to reduce our discourse to one unifying set of features.

Rush Rhees neatly expresses a similar line of thinking:

> [Wittgenstein] would constantly describe 'different ways of doing it', but he did not call them different ways of saying the same thing. Nor did he think we could reach the heart of the matter by seeing what they all have in common. He did not see them as so many fumbling attempts to say what some of them ever does say perfectly. The variety is important – not in order to fix your gaze on the unadulterated form, but to keep you from looking for it. (1965, p.40)

A proper understanding of the variety of ways in which our practices operate, including the vast amount of ways certain words are used in those practices, will lead us away from thinking that there is some uniquely identifying property or set of properties to which all uses must correspond:

> A *picture* held us captive. And we could not get outside it, for it lay in our language and language seemed to repeat it to us inexorably. (§115)

Recognition of the variety of uses of our concepts within such practices stops us from trying to provide an all-encompassing system in the first place.

<u>Non-logical privacy, and developments:</u>

If, as has been suggested, Wittgenstein was not opposed to sensation concepts untied to external circumstance, physiology, or behaviour then there is scope for retaining the original 'natural' intuition that a central feature of such concepts is how they feel *intrinsically*. I do not think we should shy away from this conclusion. In my view, one that is arguably at the least consonant with, if not emphasised by, Wittgenstein, is that it should be seen as *part of the language-game* that phenomenal concepts can be understood as concepts whose content is behaviourally, physiologically and circumstantially inaccessible to others, where 'the feeling itself' is what matters and can be conceptually divorced from any attendant extrinsicalities. This sort of privacy, in my view, is (trivially) acceptable, and recognisable as something one 'plays by oneself' (§248), and of which others ultimately cannot gain access to. Indeed, I recommend an even more radical position than this, arguing that such intuitions extend towards deeper conclusions than the relatively far-flung problems about Marianna, spectrum inversion, and zombies reveal. I want to suggest that the notion of 'privacy' discussed here plays a myriad of roles in our consciousness-ascribing practices. The *privacy* of phenomenal concepts is an important and central feature of our ordinary life. That we *don't* have immediate access to others experiences, that they are not linked to any external circumstances, etc., is to be admitted. This is not something that philosophers should be aiming to refute or downplay. It is not precluded by the community's inability to access it directly, or by a lack of connection with external behaviour. On the contrary, it is part of the 'grammar' of such a concept that its intrinsic features are, fundamentally, hidden from others.

What I mean by privacy being part of our lives is that there are some aspects of our conscious lives that are unknown or unknowable to others. There is something essentially private to our experiences. In fact, I believe examples of this form of privacy express fairly deep aspects of personhood and the knowledge we possess of the conscious lives of those around us. Moreover, as discussed in my fourth response in Part II, there is often something it is like to be a certain kind of individual such that the experiences of that individual are not understandable by others who do not possess similar sentiments, or who do not come from similar enough backgrounds, lifestyles, or temperaments (etc.). (This is indeed even further emphasised when it comes to understanding the experiences of animals, and the sort of access we have to their conscious lives (see Nagel, 1986). We might also consider it in relation to our capacities and limitations for *self*-knowledge). By implication, there is something it is like for a particular conscious being to be the kind of conscious being that it is, this something can be entirely inaccessible, and yet this does not therefore render it *logically* private.

I do not think that Wittgenstein would have been adamant that this might not be so, that we might not say such a thing. Consider a remark such as:

> "What anyone says to himself within himself is hidden from me" might of course also mean that I can for the most part not *guess* it, nor can I read it off from, for example, the movement of his throat (which would be a possibility.) (*PI, Part II,* §xi, p.188)

Moreover, we can appeal to a host of other remarks where Wittgenstein may be read as recognising a gap between pain and pain behaviour, e.g. *PI* §§278, 281, 291, 304, or between colour sensations and their external counterparts, e.g. *PI* §§276-8. Drawing from this idea, and attempting to develop it, we can go on to consider interesting cases from everyday life where this gap can be seen in operation.

My view partly turns on arguing against a trend in philosophy which assumes that the possibility of not having such knowledge of others minds represents a puzzle that philosophers should seek to overcome. My basic starting point is that such 'knowledge' is more difficult to obtain than supposed, and that there are many ways in which understanding is limited (here I agree with Nagel (see his 1986, chapter 3). The position I argue against often appeals to examples where e.g. two people look at the same object, say a red ball, and try to show how we can 'know' that the other person sees the same (or similar) as we do. Or they appeal to other such very basic sensory-perceptual experiences. From my view, of course we should accept that we see the same (or similar) colours as others. My point is that when we start considering deeper facets of personhood— when we start talking about beliefs, about world-views, about morality, politics, social responsibility etc. etc. it is far more difficult to say that we can just 'see' the same things or understand how others think and feel. It is these facets which I think constitute in a much deeper way what it is like to be conscious, much deeper than simply 'seeing a red ball'. Contrast the life of Queen Elizabeth II with a Russian street sweeper, or that between a high-flying corporate lawyer and a committed Benedictine monk. Here (as I imagine it), there may be little understanding between these characters. It isn't just that they have different opinions, but that their starting points and life experiences are so far removed that they are unlikely to understand one another in much more than a basic sense. These examples, I suggest, represent an extreme of a phenomenon which affects our understanding of others in a more commonplace way. It suggests that knowledge of others minds is far more of a matter of degree, where we can often only approximate towards understanding, and sometimes fail miserably altogether. It suggests that there are limitations, sometimes insurmountable ones, based on great differences in values, temperaments, intelligence, politics, religion, or culture. There are physiological barriers, linguistic barriers, and conceptual barriers. In short, there are many ways in which our understanding is limited than is typically supposed by the 'all or nothing' manner of speaking which dominates discussion of knowledge of other minds between sceptics and their opponents.

There is the flip-side to this, which is the extent to which people *can* understand each other. This is the more optimistic picture; with all the ways in which we can be separated mentally, culturally, emotionally etc. there are ways in which connections are made, where it makes sense to say that someone truly

understands us. People from totally different backgrounds can still forge perfectly meaningful understanding, even in the face of all these barriers; there is something to be said for recognising a shared human nature in this respect—the queen and the street sweeper 'could' form a wonderful bond. Even so, I still think that for the most part, the basic starting point and often the endpoint is failures in understanding. Perhaps that is what makes understanding such an achievement in the first place.

Admittedly, these last speculations are fairly removed from concerns that Wittgenstein directly engaged in. However, I would like to think that they are consonant with them, and represent an interesting aspect of the notion of *privacy*. Unfortunately, a proper attempt at an 'übersichtliche Darstellung' is beyond the bounds of this essay. The chief purpose is limited to hinting at how such considerations might serve as 'grammatical' reminders, describing features of our conscious lives. In this way, I think *non-logical* privacy can be shown to play an important role in a full philosophy of psychology.

At this point, we reach the terminus of the argument. If correct, the purpose has been to give a satisfactory account of phenomenal concepts from a Wittgensteinian perspective. During the course of which, I have been critical of much of the exegetical work to date, and seem to have put myself in opposition to many popularly defended lines of interpretation. I can only say that this is how I read Wittgenstein. One might say that at best I have provided reasons *not* to be a Wittgensteinian. Perhaps this is true. On the other hand, with more optimism, I like to think that I have defended a line of interpretation that accommodates for some of the more interesting puzzles that have been arising in philosophy of mind, especially during the last ten to twenty years. And with even more optimism, I have provided a way of thinking about the PLA which approximates to Wittgenstein's true intentions, and gives pause for thought about ways in which we might develop those insights in interesting new ways.

BIBLIOGRAPHY:

Wittgenstein's works, written and recorded:

AWL (1979) *Wittgenstein's Lectures, Cambridge 1932-1935*, from the notes of Ambrose, A. and MacDonald, M., Ambrose, A. (ed.), Oxford: Blackwell

CV (1980) *Culture and value,* rev. ed., Winch, P. (trans.), Oxford: Blackwell

LCA (1966) *Lectures and conversations on aesthetics, psychology and religious belief*, Barrett, C. (ed.), Oxford: Blackwell

NLP	(1968) 'Notes for Lectures on "Private Experience" and "Sense Data", *The Philosophical Review*, 77 (3), pp.333-368
OC	(1969) *On Certainty*, Anscombe, G. & von Wright, G. (edd.), New York: Harper
PI	(2001) *Philosophical Investigations*, 3rd edition, Anscombe, G.E.M. (trans.), Oxford: Blackwell
PO	(1993) *Philosophical Occasions 1912-1951*, Klagge, J. & Nordmann, A. (edd.), Indianapolis: Hackett
PR	(1975) *Philosophical Remarks*, Rhees, R. (ed.), Hargreaves, R. & White, R. (trans.), Oxford: Blackwell
RC	(1977), *Remarks on Colour*, McAlister, L. & Schättle, M. (trans.), Oxford: Blackwell
RFM	(1978) *Remarks on the Foundations of Mathematics*, 3rd edition, von Wright, G., Rhees, R., & Anscombe, G.E.M. (edd.), Anscombe, G.E.M. (trans.), Oxford: Blackwell
RPP, Vol. 2	(1980) *Remarks on the Philosophy of Psychology, Volume 2*, von Wright, G. & Nyman, H., (edd.), Chicago: University of Chicago Press

---

Ayer, A.: 1986, *Wittgenstein*, London: Penguin

Bar-On, D: 2004, *Speaking my Mind*, Oxford: Oxford University Press

Ball, D.: 2009, 'There are no phenomenal concepts', *Mind*, Vol.118, pp.935-962.

Balog, K.: 2009, 'Phenomenal Concepts', in McLaughlin, B. et al (edd.), *The Oxford Handbook of Philosophy of Mind*, Oxford: Oxford University Press.

Block, N.: 2007, 'Wittgenstein and Qualia', *Philosophical Perspectives, Vol. 21*, pp.73-115.

Canfield, J.: 1996, 'The Community View', *The Philosophical Review*, Vol.105, pp.469-488

Chalmers, D.: 1996, *The Conscious Mind*, Oxford: Oxford University Press

Chalmers, D.: 2003, 'The Content and Epistemology of Phenomenal Belief', in Smith, Q. & Jokic, A. (edd.), *Consciousness: New Philosophical Perspectives,* Oxford: Oxford University Press.

Chalmers, D.: 2010, *The Character of Consciousness,* Oxford: Oxford University Press.

Child, W.: 2011, *Wittgenstein*, (London: Routledge).

Child, W.: forthcoming, 'Wittgenstein, phenomenal concepts and what it's like'.

Churchland, P.: 1986, *Neurophilosophy*, Cambridge, MA: MIT Press.

Cottingham, J.: 2010, 'Wittgenstein, religion, and analytic philosophy', in H.-J. Glock, H-J, and Hyman, J. (edd.), *Wittgenstein and analytic philosophy, essays for P.M.S. Hacker*, Oxford: Oxford University Press.

Crane, T., 2000, 'The origins of qualia', in Crane, T. and Patterson, S. (edd.), *The History of the Mind-Body Problem*, London: London Studies in the History of Philosophy.

Day, W.: 2010, 'Wanting to Say Something: Aspect-Blindness and Language', in Day, W. & Krebs, V. (edd.), *Seeing Wittgenstein Anew*, Cambridge: Cambridge University Press.

Dennett, D.: 2007, 'What RoboMary knows', in Alter, T. (ed.), *Phenomenal Concepts and Phenomenal Knowledge*, Oxford: Oxford University Press.

Fogelin, R.: 1987, *Wittgenstein*, second edition, London: Routledge.

Fogelin, R.: 2009, *Taking Wittgenstein at his word: A textual study*, Princeton: Princeton University Press.

Ginet, C.: 1968, 'How words mean kinds of sensations', *Philosophical Review,* 77.

Glock, H-J.: 1996, *A Wittgenstein Dictionary*, Oxford: Blackwell.

Hacker, P.M.S.: 1986, *Insight and Illusion: Themes in the Philosophy of Wittgenstein,* Revised 2nd edition, Oxford: Oxford University Press.

Hacker, P.M.S.: 1993a, *Wittgenstein: Meaning and Mind*, Vol. 3. of *An Analytical Commentary on the* Philosophical Investigations. *Part I — Essays*, Oxford: Blackwell.

Hacker, P.M.S.: 1993b, *Wittgenstein: Meaning and Mind*, Vol. 3. of *An Analytical Commentary on the* Philosophical Investigations. *Part II — Exegesis*, Oxford: Blackwell.

Hacker, P.M.S.: 2005, *Wittgenstein: Understanding and Meaning*, Vol. 1. of *An Analytical Commentary on the* Philosophical Investigations. *Part I – Essays*, revised edition*,* Oxford: Blackwell.

Hacker, P.M.S.: 2002, 'Is there anything it is like to be a bat?', *Philosophy*, 77, pp.157-174.

Hacker, P.M.S.: 2012, 'The sad and sorry history of consciousness: being among other things a challenge to the "consciousness studies community"', *Institute of Philosophy, supplementary volume 70*, pp.149-168.

Hanfling, O.: 2001a, 'Consciousness: 'The last mystery", in Schroeder, S., (ed.), *Wittgenstein and Contemporary Philosophy of Mind*, Basingstoke: Palgrave, pp.36-59.

Hanfling, O.: 2001b, 'Thinking', in Glock, H-J., *Wittgenstein: A Critical Reader*, Oxford: Blackwell.

Hopkins, J.: 1974, 'Wittgenstein and Physicalism', *Proceedings of the Aristotelian Society* 75, pp.121-146.

Horwich, P.: 2012, *Wittgenstein's Metaphilosophy,* Oxford: Oxford University Press.

Jackson, F.: 1982, 'Epiphenomenal Qualia', *Philosophical Quarterly,* 32, pp.127-36.

Kilverstein, J.: 2008, 'Wittgenstein, Qualia and the Autonomy of Grammar', in Levy, D. & Zamuner, E. (edd.), *Wittgenstein's Enduring Arguments*, London: Routledge

Kirk, R.: 1999 'Why there couldn't be zombies', *Proceedings of the Aristotelian Society, Supplementary Vol. LXIII*, pp.1-16

Kripke, S.: 1979, 'A puzzle about belief' in *Meaning and Use*, Margalit, A., (ed.), Dordrecht: D. Reidel.

Kripke, S.: 1982, *Wittgenstein on Rules and Private Language*, Cambridge, Mass.: Harvard University Press.

Kusch, M.: 2006, *A Sceptical Guide to Meaning and Rules: Defending Kripke's Wittgenstein*, Chesham: Acumen.

Kuusela, O.: 2008, *The Struggle Against Dogmatism: Wittgenstein and the Concept of Philosophy*, Cambridge, Mass.: Harvard University Press.

Loar, B.: 1990/1997, 'Phenomenal States', original version in *Philosophical Perspectives 4: Action Theory and Philosophy of Mind*, pp.81-108, revised version in Block, N. et al. (edd.), *The Nature of Consciousness*, Cambridge, MA: MIT Press.

Levine, J.: 1983, 'Materialism and qualia: the explanatory gap', *Pacific Philosophical Quarterly* 64, pp.354-61.

Lycan, W.: 1990, 'What is the Subjectivity of the Mental', in *Philosophical Perspectives* 4: *Action Theory and Philosophy of Mind,* pp.109–130.

McDowell, J.: 1984, 'Wittgenstein on Following a Rule', *Synthese*, 58, pp.325-63.

McGinn, C: 1984, *Wittgenstein on Meaning*, Oxford: Blackwell.

McGinn, M.: 1997, *Routledge Philosophy Guidebook to Wittgenstein and the Philosophical Investigations*', London: Routledge.

Malcolm, N.: 1984a, *Ludwig Wittgenstein: A Memoir,* Oxford: Oxford University Press.

Malcolm, N.: 1984b, 'Consciousness and causality' in Armstrong and Malcolm, 1984, *Consciousness and Causality* Oxford: Blackwell, pp. 1-101.

Malcolm, N.: 1986, *Nothing is Hidden: Wittgenstein's Criticism of his Early Thought,* Second Edition, Oxford: Blackwell.

Malcolm, N.: 1989, 'Wittgenstein on Language and Rules', *Philosophy*, Vol.64, pp.5-28.

Miller, A. & Wright, C.: 2002, *Rule-Following and Meaning*, Durham: Acumen.

Millikan, R.: 1984, *Language, Thought, and Other Biological Categories*, Cambridge, Mass.: MIT Press.

Monk, R.: 1990, *Ludwig Wittgenstein: The Duty of Genius*, London: Vintage Books.

Nagel, T.: 1974, 'What is it like to be a bat?', *Philosophical Review* 83, pp.435-50.

Nagel, T,: 1986, *The View from Nowhere*, 1986, Oxford: Oxford University Press.

Nemirow, L.: 2006, 'So This is What it's Like: a Defense of the Ability Hypothesis', in Alter, T. & Walter, S., (edd.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, Oxford: Oxford University Press.

Nida-Rümelin, M.: 1996, 'What Mary couldn't know', in Metzinger, T. (ed.), *Conscious Experience*, Schoenigh: Paderborn.

Papineau, D.: 1999, 'Normativity and judgement'*, Aristotelian Society Supplementary Volume*, 73, pp.17-43.

Papineau, D.: 2002, *Thinking about Consciousness*, Oxford: Oxford University Press.

Papineau, D.: 2006, 'Phenomenal and perceptual concepts', in T. Alter & S. Walter (edd.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, Oxford: Oxford University Press.

Papineau, D.: 2011, 'Phenomenal concepts and the Private Language Argument', *American Philosophical Quarterly,* Vol.48, pp.175-184.

Peacocke, C.: 1992, *A Study of Concepts*, Cambridge, Mass.: MIT Press.

Peacocke, C.: 2009, *Truly Understood*, Oxford: Oxford University Press.

Pears, D.: 1988, *The False Prison: A Study of the Development of Wittgenstein's Philosophy*, Vol.2, Oxford: Clarendon Press.

Pears, D.: 2006, *Paradox and Platitude in Wittgenstein's Philosophy*, Oxford: Oxford University Press.

Perry, J.: 2001, *Knowledge, Possibility, and Consciousness*, Cambridge, Mass.: MIT Press.

Pitcher, G.: 1965, 'Wittgenstein, nonsense, and Lewis Carroll', *The Massachusetts Review*, Vol. 6, pp.591-611.

Putnam, H.: 1968, 'Brains and Behaviour', in Butler, R. (ed.), *Analytical Philosophy, Second series*, Oxford: Blackwell, pp.1-19.

Rhees, R.: 1965, 'Some Developments in Wittgenstein's View of Ethics', 1965, reprinted in Rhees, R., *Moral Questions*, 1999, Basingstoke: Macmillan.

Russell, B.: 1912, *The Problems of Philosophy,* Oxford: Oxford University Press.

Schroeder, S.: 2006, *Wittgenstein*, Cambridge: Polity Press.

Schroeder, S: 2007, 'The tightrope walker', *Ratio,* Vol. 20, pp.442-463.

Shoemaker, S.: 1982, 'The Inverted Spectrum', reprinted in *Identity, Cause, and Mind*, 1984, Cambridge: Cambridge University Press, pp.357-81.

Stern, D.: 1994, 'Review Essay: Recent Work on Wittgenstein, 1980-1990)', *Synthese*, Vol.98, pp.415-458.

Stern, D.: 2004, Wittgenstein's *Philosophical Investigations*: An Introduction, Cambridge: Cambridge University Press.

Stern, D.: 2010, 'Wittgenstein, qualia and the inverted spectrum', in *Wittgenstein: Certeza?* Moreno, A. (ed.), Campinas, Brazil: UNICAMP, pp 49-63.

Stich, S.: 1983, *From Folk Psychology to Cognitive Science*, Cambridge, MA: MIT Press.

Stoljar, D.: 2005, 'Physicalism and Phenomenal Concepts', *Mind and Language*, vol.20, pp.469-494.

Sturgeon, S.: 1994, 'The epistemic view of subjectivity', *The Journal of Philosophy,* vol.91, pp.221-235.

Tanney, J.: 2004, 'On the Conceptual, Psychological, and Moral Status of Zombies, Swamp- Beings, and Other 'Behaviourally Indistinguishable' Creatures', *Philosophy and Phenomenological Research Vol. LXIX, No. 1*.

Tripodi, P.: 2013, 'Wittgenstein on the Gulf between Believers and Non-believers', *Philosophia, 41,* pp.63-79.

Tye, M.: 1995, *Ten Problems of Consciousness*, Cambridge, MA: MIT Press.

Tye, M.: 2000, *'Knowing What it is Like: The Ability Hypothesis and the Knowledge Argument'* in Tye, M., *Consciousness, Color, and Content*, Cambridge, MA: MIT Press.

Tye, M.: 2003, '*A Theory of Phenomenal Concepts',* in O' Hear, A. (ed.), *Minds and Persons*, Cambridge: Cambridge University Press.

Tye, M.: 2009, *Concsiousness Revisited*, Cambridge, Mass.: Bradford Book, MIT Press.

Wright, C.: 1998, 'Self-knowledge: The Wittgensteinian legacy', in Wright, Smith, and MacDonald (edd.), *Knowing our Own Minds*, Oxford: Oxford University Press.